

GenCore version 5.1.6 Copyright (c) 1993 - 2004 CompuGen Ltd.									
OM protein - protein search, using sw model									
Run on: May 11, 2004, 13:23:11 ; Search time 47 Seconds (without alignments) 120.233 Million cell updates/sec									
Title: US-09-171-432A-38									
Perfect score: 109									
Sequence: 1 VDDPRSEEDKRFESHIECRK 20									
Scoring table: BLOSUM62									
Gapop 10.0 , Gapext 0.5									
Searched: 1586107 seqs, 282547505 residues									
Total number of hits satisfying chosen parameters: 1586107									
Minimum DB seq length: 0									
Maximum DB seq length: 2000000000									
Post-processing: Minimum Match 0%									
Maximum Match 100%									
Listing first 100 summaries									
Database : A_Geneseq_29Jan04.*									
1: geneseqp1980s.*									
2: geneseqp1990s.*									
3: geneseqp2000s.*									
4: geneseqp2001s.*									
5: geneseqp2002s.*									
6: geneseqp2003as.*									
7: geneseqp2003bs.*									
8: geneseqp2004s.*									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description				
1	109	100.0	20	AAW42921	Immunogen				
2	109	100.0	21	AAW42921	Synthetic				
3	109	100.0	854	AAW50287	Hepatitis				
4	109	100.0	1077	AAW95559	A partial				
5	109	100.0	1091	AAW32426	Translate				
6	109	100.0	2227	AAW05697	Attenuate				
7	109	100.0	2227	AAW34074	Hepatitis				
8	109	100.0	2227	AAW18609	Amino aci				
9	109	100.0	2227	AAW18607	Amino aci				
10	109	100.0	2227	AAW18608	Amino aci				
11	109	100.0	2227	ABG31729	Attenuate				
12	109	100.0	2227	ABG31727	Wild-type				
13	109	100.0	2227	ABG31728	Hepatitis				
14	109	100.0	2227	ABU08640	Attenuat				
15	109	100.0	2227	ABU08641	Attenuat				
16	109	100.0	2227	ABU08639	Wild type				
17	106	97.2	366	AAW50230	Sequence				
18	106	97.2	993	AAW50116	Sequence				
19	106	97.2	993	AAW50231	Sequence				
20	106	97.2	2227	AAE19899	Hepatitis				
21	106	97.2	2227	ABW00350	Hepatitis				
22	87	79.8	2227	AAW60066	Sequence				
23	81	74.3	839	AAW15629	Capsid re				
24	69	63.3	20	AAW42920	Immunogen				
25	69	63.3	21	AAW42921	Synthetic				

26	69	63.3	300	1	AAW42921	Portion o
27	50	45.9	409	6	AAW42921	Photorhab
28	46.5	42.7	183	3	AAW42921	Arabidops
29	46.5	42.7	187	3	AAW42921	Arabidops
30	46.5	42.7	200	3	AAW42921	Arabidops
31	46.5	42.7	200	3	AAW42921	Arabidops
32	46	42.2	306	4	AAW42921	Novel hum
33	46	42.2	701	4	AAW42921	Mouse hum
34	44.5	40.8	1334	2	AAW42921	Protein e
35	44	40.4	131	3	AAW42921	Arabidops
36	44	40.4	134	3	AAW42921	Arabidops
37	44	40.4	231	2	AAW42921	Chlamydia
38	44	40.4	285	4	AAW42921	Human pro
39	44	40.4	292	4	AAW42921	Human pro
40	44	40.4	326	3	AAW42921	Arabidops
41	44	40.4	383	3	AAW42921	Arabidops
42	44	40.4	498	3	AAW42921	Arabidops
43	43	39.9	598	5	AAW42921	Neisseria
44	43	39.9	818	6	AAW42921	Acinetoba
45	43	39.4	125	3	AAW42921	Human sec
46	43	39.4	125	4	AAW42921	Novel hum
47	43	39.4	242	6	AAW42921	Photorhab
48	43	39.4	292	3	AAW42921	Arabidops
49	43	39.4	293	3	AAW42921	Arabidops
50	43	39.4	322	3	AAW42921	Arabidops
51	43	39.4	325	3	AAW42921	Arabidops
52	43	39.4	325	3	AAW42921	Arabidops
53	43	39.4	379	3	AAW42921	Arabidops
54	43	39.4	382	3	AAW42921	Arabidops
55	43	39.4	382	3	AAW42921	Arabidops
56	43	39.4	429	3	AAW42921	Arabidops
57	43	39.4	494	3	AAW42921	Arabidops
58	43	39.4	497	3	AAW42921	Arabidops
59	43	39.4	497	6	AAW42921	Arabidops
60	43	39.4	502	3	AAW42921	Arabidops
61	43	39.4	505	3	AAW42921	Arabidops
62	43	39.4	710	5	AAW42921	Human pol
63	43	39.4	780	2	AAW42921	Mitochond
64	43	39.4	780	4	AAW42921	Amino aci
65	43	39.4	780	7	AAW42921	Human src
66	43	39.4	802	4	AAW42921	Novel hum
67	43	39.4	814	4	AAW42921	Novel hum
68	43	39.4	1606	6	AAW42921	Human DIT
69	43	39.4	1715	4	AAW42921	Drosophil
70	43	39.4	2141	6	AAW42921	Human DIT
71	43	39.4	19318	6	AAW42921	Streptomy
72	43	39.4	405	7	AAW42921	Immunogen
73	42.5	39.0	20	2	AAW42921	Synthetic
74	42	38.5	21	4	AAW42921	Alloioococ
75	42	38.5	133	6	AAW42921	Human bre
76	42	38.5	145	4	AAW42921	Antigen t
77	42	38.5	180	2	AAW42921	Arabidops
78	42	38.5	187	3	AAW42921	Human ORF
79	42	38.5	256	3	AAW42921	Human pro
80	42	38.5	398	4	AAW42921	Human pro
81	42	38.5	453	4	AAW42921	Arabidops
82	42	38.5	474	3	AAW42921	Arabidops
83	42	38.5	477	3	AAW42921	Arabidops
84	42	38.5	533	3	AAW42921	Human com
85	42	38.5	560	2	AAW42921	Drosophil
86	42	38.5	661	4	AAW42921	Protein s
87	42	38.5	679	6	AAW42921	Drosophil
88	42	38.5	704	4	AAW42921	Protein e
89	42	38.5	704	6	AAW42921	Drosophil
90	42	38.5	722	5	AAW42921	Human SUL
91	42	38.5	786	2	AAW42921	Human/Mur
92	42	38.5	1001	5	AAW42921	Novel hum
93	42	38.5	1001	5	AAW42921	Human pro
94	42	38.5	1001	7	AAW42921	Breast ca
95	42	38.5	1005	6	AAW42921	Human TAO
96	42	38.5	1166	7	AAW42921	Human pro

99 42 38.5 2759 6 AAO16418 Human nuc
100 42 38.5 2781 3 AAY57453 Human tra

ALIGNMENTS

RESULT 1
AAW42921
ID AAW42921 standard; peptide; 20 AA.
AC AAW42921;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1314.
XX
KW Immunogenic peptide; immunogenic epitope; VP1 protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN W05740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
PS Claim 13; Page 111; 140pp; English.
XX
CC Peptides AAW42906-21 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the VP1
CC protein of HAV corresponding to amino acids 492-791. The present peptide
CC is derived from amino acids 779-798, and has a reactivity of 54.2% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 109; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 1 VDDPRSEEDKRFESHIECRK 20

RESULT 2
AAB69438
ID AAB69438 standard; peptide; 21 AA.
XX
AC AAB69438;
XX
DT 20-APR-2001 (first entry)
XX

DE Synthetic HAV VP1 peptide, SEQ ID NO: 38.
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN W0200105824-A2.
XX
PD 25-JAN-2001.
XX
PF 14-JUL-2000; 2000WO-US019267.
XX
PR 15-JUL-1999; 99US-0144412P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
WPI; 2001-112681/12.
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
XX
PS Claim 10; Page 92; 130pp; English.
XX

CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC Igm antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or biofluorescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the Igm antibody reactivity

XX Sequence 21 AA;
SQ
Query Match 100.0%; Score 109; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 1 VDDPRSEEDKRFESHIECRK 20

RESULT 3
AAP50287
ID AAP50287 standard; protein; 854 AA.
XX
AC AAP50287;
XX
DT 25-MAR-2003 (revised)
DT 30-NOV-1991 (first entry)
XX
DE Hepatitis A virus (HAV) peptide corresponding to the capsid protein
DE region of poliovirus RNA.
XX
KW Hepatitis A virus assay; antigen; antibody.
XX
OS Hepatitis A virus.
XX
PN W08501517-A.
XX

PD 11-APR-1985.
 XX
 PF 27-SEP-1984; 84WO-US001552.
 XX
 PR 30-SEP-1983; 83US-00537911.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racaniello VR;
 XX
 DR WPI; 1985-098846/16.
 DR N-PSDB; AAN50330.
 XX
 PT New hepatitis A virus cDNA - useful in assays for the virus and for
 PT prodn. of the viral antigen and antibodies to it.
 XX
 PS Example; Fig 7; 60pp; English.
 XX
 CC The inventors claim HAV cDNA and a method for producing it, whereby large
 CC ams. can be obt'd. economically. The cDNA is useful in the assay for
 CC detection of HAV quickly and easily and with high sensitivity and
 CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
 CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 XX Sequence 854 AA;
 SQ
 Query Match 100.0%; Score 109; DB 1; Length 854;
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDKRFESHIECRK 20
 DB 779 VDDPRSEEDKRFESHIECRK 798
 RESULT 4
 ID AAW95559 standard; protein; 1077 AA.
 AC AAW95559;
 DT 28-APR-1999 (first entry)
 DE A partial hepatitis A virus (HAV) protein.
 KW Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX
 OS Hepatitis A virus.
 PN US5849562-A.
 PD 15-DEC-1998.
 XX
 PF 06-JUN-1995; 95US-00468926.
 XX
 PR 30-SEP-1983; 83US-00537911.
 PR 27-SEP-1984; 84US-00654942.
 PR 06-OCT-1988; 88US-00256135.
 PR 06-NOV-1991; 91US-00788262.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Emerson SU, Purcell RH;
 DR WPI; 1999-094412/08.
 DR N-PSDB; AAX01006.
 XX
 PT Chimeric hepatitis A virus strains - with P2 region from cell-culture-
 PT adapted strain in wild-type genome.
 XX

PS Disclosure; Fig 7A-L; 36pp; English.
 XX
 CC The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting of a wild
 CC type HAV genome in which the P2 region is replaced by the P2 region from
 CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
 CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
 CC sufficient for establishment of infection and accelerated growth in cell
 CC culture
 XX
 SQ Sequence 1077 AA;
 Query Match 100.0%; Score 109; DB 2; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 7.6e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDKRFESHIECRK 20
 DB 1002 VDDPRSEEDKRFESHIECRK 1021
 RESULT 5
 ID AAR32426 standard; protein; 1091 AA.
 AC AAR32426;
 XX
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-DEC-2001 (revised)
 DT 10-JUN-1993 (first entry)
 XX
 DE Translated from 5' region of Hepatitis A Virus genomic clone.
 XX
 KW HAV HM-175; chronic liver disease; picornavirus.
 XX
 OS Hepatitis A virus.
 XX
 FH Key Location/Qualifiers
 FT Region 1..711
 FT /note= 'x's correspond to nonsense codons, i.e. this
 FT region is not an ORF"
 FT Region 238..1091
 FT /label= ORF
 FT /note= "second putative initiation codon at position 240"
 XX
 XX USN7788262-N.
 XX
 PD 15-DEC-1992.
 XX
 PF 06-NOV-1991; 91US-00788262.
 XX
 PR 30-SEP-1983; 88US-00536911.
 PR 27-SEP-1984; 84US-00654942.
 PR 06-OCT-1988; 88US-00256135.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racaniello VR, Baroudy BM, Emerson SU;
 XX
 DR WPI; 1993-067429/08.
 DR N-PSDB; AAQ36934.
 XX
 PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
 PT antigen and antibodies.
 XX
 PS Disclosure; Fig 7; 65pp; English.
 XX
 CC HAV virion RNA was extracted from the livers of marmosets which had been
 CC inoculated with HAV (the HAV had previously been passaged twice in
 CC marmosets). The RNA was used to prepare cDNA clones by standard
 CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected

CC African Green Monkey kidney cells were selected for further analysis. A
CC 7.4kb restriction map (about 99% of the HAV genome) was constructed from
CC 5 overlapping inserts. The sequence of the first 3.3kb (approx.) from the
CC 5'-terminus was determined. An amino acid sequence was deduced from the
CC entire clone and an open reading frame was identified starting at
CC position 238. A comparison of the predicted HAV amino acid sequences with
CC the known capsid protein sequences of other picornaviruses (poliovirus,
CC foot and mouth disease virus and encephalomyelitis virus) revealed areas
CC of local homology. (Note: Revised entry submitted to correct the patent
CC number format of US Government-owned NTIS applications to prevent clashes
CC with ongoing US granted patent numbers. For further information please
CC visit the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html.)
CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 1091 AA;

Query Match 100.0%; Score 109; DB 2; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||
Db 1016 VDDPRSEEDKRFESHIECRK 1035

RESULT 6
AA05697
ID AAR05697 standard; protein; 2227 AA.

AC AAR05697;
XX
DT 24-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 15-AUG-1990 (first entry)
XX
DE Attenuated hepatitis A virus.
XX
KW Hepatitis A virus; vaccine; attenuated.
XX
OS Hepatitis A virus; strain HM-175.

Key	Location/Qualifiers
FT Region	1..23
FT Region	/label= VP4 = 1A
FT Region	24..245
FT Region	/label= VP2 = 1B
FT Region	246..491
FT Region	/label= VP3 = 1C
FT Region	492..791
FT Region	/label= VP1 = 1D
FT Region	792..980
FT Region	/label= 2A
FT Region	981..1087
FT Region	/label= 2B
FT Region	1088..1422
FT Region	/label= 2C
FT Region	1423..1496
FT Region	/label= 3A
FT Region	1497..1519
FT Region	/label= 3B = VPg
FT Region	1520..1738
FT Region	/label= 3C
FT Region	1739..2227
FT Region	/label= 3D

XX US4894228-A.
PN
XX
PD 16-JAN-1990.
XX
PF 12-JUL-1988; 88US-00217824.
XX

PR 19-SEP-1984; 84US-00652067.
PR 09-SEP-1986; 86US-00905146.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
PI Daemer RJ, Gust ID;
XX
XX WPI; 1990-075557/10.
DR N-PSDB; AAQ03512.
XX
XX Vaccine against hepatitis A virus infection - comprises novel attenuated
FT hepatitis A virus strain.
XX
XX Claim 1; Fig 1; 18pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 7
AAW34074
ID AAW34074 standard; protein; 2227 AA.

AC AAW34074;
XX
DT 17-OCT-2003 (revised)
DT 27-APR-1998 (first entry)
XX
DE Hepatitis A virus HM-175 protein sequence.
XX
KW HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
KW vaccine
XX
OS Hepatitis A virus; HM-175.

Key	Location/Qualifiers
FT Protein	1..23
FT Protein	/label= VP4
FT Protein	24..245
FT Protein	/label= VP2
FT Protein	246..491
FT Protein	/label= VP3
FT Protein	492..791
FT Protein	/label= VP1
FT Protein	792..980
FT Protein	/label= 2A
FT Protein	981..1087
FT Protein	/label= 2B
FT Protein	1088..1422
FT Protein	/label= 2C
FT Protein	1423..1496
FT Protein	/label= 3A
FT Protein	1497..1519
FT Protein	/label= 3B
FT Protein	1520..1738

```

FT      /label= 3C
FT      1739..2227
FT      /label= 3D
XX      WO9740166-A2.
XX      30-OCT-1997.
XX      18-APR-1997; 97WO-US006506.
XX      19-APR-1996; 96US-0015642P.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      Raychaudhuri G, Emerson SU, Purcell RH,
XX      WPI; 1997-535850/49.
XX      N-PSDB; AAT93023.
XX      Human attenuated HAV genome containing simian HAV 2C gene - useful as
XX      vaccines against HAV infection.
XX      Disclosure; Fig 13A-D; 66pp; English.
XX      This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
XX      175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
XX      by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
XX      construct (i) comprises a genome of HAV, where the genome is a human
XX      attenuated HAV genome in which a region of the 2C gene has been replaced
XX      by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
XX      (see AAT93024). The region of the 2C gene from AGM-27 contained in the
XX      construct preferably encodes amino acids 120-328 of the 2C protein, amino
XX      acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
XX      of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3)
XX      a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
XX      cell containing the HAV of (3). (1) or its RNA transcript, can be used as
XX      a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can
XX      also be used to stimulate the production of protective antibodies in the
XX      mammal. (Updated on 17-OCT-2003 to standardise OS field)
XX      Sequence 2227 AA;
XX      Query Match 100.0%; Score 109; DB 2; Length 2227;
XX      Best Local Similarity 100.0%; Pred. No. 1.7e-07;
XX      Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 VDDPRSEEDKRFESHIECRK 20
DB      779 VDDPRSEEDKRFESHIECRK 798
XX      RESULT 8
XX      AAB18609
XX      ID AAB18609 standard; protein; 2227 AA.
XX      AC AAB18609;
XX      DT 15-JAN-2001 (first entry)
XX      Amino acid sequence of live attenuated Hepatitis A virus 4380.
XX      HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX      HAV 4380.
XX      Hepatitis A virus.
XX      US6113912-A.
XX      05-SEP-2000.
XX      07-JUN-1995; 95US-00475886.
XX      Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX      line useful as vaccine for protecting humans against hepatitis A virus
XX      infection, has modified genome compared to wild type.
XX      Disclosure; Fig 6A-K; 72pp; English.
XX      The present sequence is derived from a wild type hepatitis A virus (HAV)

```

```

PR      17-SEP-1993; 93WO-US008610.
PR      17-APR-1995; 95US-00397232.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX      WPI; 2000-586464/55.
XX      N-PSDB; AAA75478.
XX      Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX      line useful as vaccine for protecting humans against hepatitis A virus
XX      infection, has modified genome compared to wild type.
XX      Disclosure; Col 93-104; 72pp; English.
XX      The present sequence is derived from a live attenuated hepatitis A virus
XX      (HAV) of the invention, designated HAV 4380. The sequence is produced by
XX      modifying wild type HAV strain HM-174. The HAV of the invention are
XX      adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
XX      is able to propagate in MRC-5 cells and retain appropriate attenuation.
XX      It is useful as a live vaccine for prophylaxis of hepatitis A in humans
XX      and other primates
XX      Sequence 2227 AA;
XX      Query Match 100.0%; Score 109; DB 3; Length 2227;
XX      Best Local Similarity 100.0%; Pred. No. 1.7e-07;
XX      Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 VDDPRSEEDKRFESHIECRK 20
DB      779 VDDPRSEEDKRFESHIECRK 798
XX      RESULT 9
XX      AAB18607
XX      ID AAB18607 standard; protein; 2227 AA.
XX      AC AAB18607;
XX      DT 15-JAN-2001 (first entry)
XX      Amino acid sequence of wild type Hepatitis A virus strain HM-175.
XX      HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
XX      Hepatitis A virus.
XX      US6113912-A.
XX      05-SEP-2000.
XX      07-JUN-1995; 95US-00475886.
XX      18-SEP-1992; 92US-00947338.
XX      17-SEP-1993; 93WO-US008610.
XX      17-APR-1995; 95US-00397232.
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX      WPI; 2000-586464/55.
XX      N-PSDB; AAA75476.
XX      Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX      line useful as vaccine for protecting humans against hepatitis A virus
XX      infection, has modified genome compared to wild type.
XX      Disclosure; Fig 6A-K; 72pp; English.
XX      The present sequence is derived from a wild type hepatitis A virus (HAV)

```

CC strain HM-174. The sequence is modified to produce HAV which are adapted
 CC to growth in the human fibroblast-like cell line MRC-5. The HAV is able
 CC to propagate in MRC-5 cells and retain appropriate attenuation. It is
 CC useful as a live vaccine for prophylaxis of hepatitis A in humans and
 CC other primates
 XX
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
 |||||
 Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 10
 AAB18608
 ID AAB18608 standard; protein; 2227 AA.

XX AAB18608;
 AC
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 XX P-35 virus.

XX Hepatitis A virus.

XX US6113912-A.

XX 05-SEP-2000.

XX 07-JUN-1995; 95US-00475886.

XX 18-SEP-1992; 92US-00947338.

XX 17-SEP-1993; 93WO-US008610.

XX 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

XX N-PSDB; AAA75477.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.

XX Disclosure; Col 67-78; 72pp; English.

XX The present sequence is derived from passage 35 of a wild type hepatitis
 CC A virus (HAV) strain HM-174; the resulting virus is designated P-35
 CC virus. The sequence is modified to produce HAV which are adapted to
 CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
 CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
 CC as a live vaccine for prophylaxis of hepatitis A in humans and other
 CC primates
 XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
 |||||
 Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 11
 ABG31729
 ID ABG31729 standard; protein; 2227 AA.

XX ABG31729;

XX 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)

XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.

XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 KW HAV 4380.

XX Hepatitis A virus; strain HM-175.

XX US6423318-B1.

XX 23-JUL-2002.

XX 31-AUG-2000; 2000US-00653499.

XX 17-SEP-1993; 93WO-US008610.

XX 17-APR-1995; 95US-00397232.

XX 07-JUN-1995; 95US-00475886.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX WPI; 2002-680946/73.

XX N-PSDB; ABS52789.

XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.

XX Disclosure; Col 93-104; 71pp; English.

XX The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents an attenuated
 CC hepatitis A virus 4380 polypeptide. (Updated on 29-AUG-2003 to
 CC standardise OS field)

XX Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
 |||||
 Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 12
 ABG31727

ID ABG31727 standard; protein; 2227 AA.

XX ABG31727;

XX 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)

XX Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

XX Hepatitis A virus; strain HM-175.

```

PN US6423318-B1.
XX
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-00653499.
XX
XX 17-SEP-1993; 93MO-US008610.
XX
XX 17-APR-1995; 95US-00397232.
XX
XX 07-JUN-1995; 95US-00475886.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2002-680946/73.
XX
XX N-PSDB; ABS52787.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX Disclosure; Fig 6; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents a hepatitis A virus
XX strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
XX field)
XX
XX SQ Sequence 2227 AA;
XX
XX Query Match 100.0%; Score 109; DB 5; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VDDPRSEEDKRFESHIECRK 20
XX |||||
XX Db 779 VDDPRSEEDKRFESHIECRK 798
XX
XX RESULT 13
XX ABG31728
XX ID ABG31728 standard; protein; 2227 AA.
XX
XX AC ABG31728;
XX
XX XX 29-NOV-2002 (first entry)
XX
XX DE Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
XX KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
XX virucide; mutant; pHAV/7; mutein.
XX
XX OS Hepatitis A virus; strain HM-175.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX
XX FT Misc-difference 764
XX FT /note= "Wild-type Glu substituted by Val"
XX FT Misc-difference 821
XX FT /note= "Wild-type Asn substituted by Ser"
XX FT Misc-difference 963
XX FT /label= "Wild-type Lys substituted by Arg"
XX FT Misc-difference 1052
XX FT /note= "Wild-type Ala substituted by Val"
XX FT Misc-difference 1062
XX FT /note= "Wild-type Gly substituted by Ala"
XX FT Misc-difference 1118
XX FT /note= "Wild-type Lys substituted by Met"
XX FT Misc-difference 1151
XX FT /note= "Wild-type Glu substituted by Lys"
XX FT Misc-difference 1163
XX
XX FT /note= "Wild-type Phe substituted by Ser"
XX FT Misc-difference 1277
XX FT /note= "Wild-type Val substituted by Ile"
XX FT Misc-difference 1500
XX FT /note= "Wild-type His substituted by Tyr"
XX FT Misc-difference 1805
XX FT /note= "Wild-type Asp substituted by Asn"
XX FT Misc-difference 1930
XX FT /note= "Wild-type Ser substituted by Thr"
XX
XX PN US6423318-B1.
XX
XX XX 23-JUL-2002.
XX
XX XX 31-AUG-2000; 2000US-00653499.
XX
XX XX 17-SEP-1993; 93MO-US008610.
XX
XX XX 17-APR-1995; 95US-00397232.
XX
XX XX 07-JUN-1995; 95US-00475886.
XX
XX XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX XX WPI; 2002-680946/73.
XX
XX XX N-PSDB; ABS52788.
XX
XX XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX
XX XX Example 3; Col 67-78; 71pp; English.
XX
XX XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents a hepatitis A virus
XX mutant strain HM-175/7 (pHAV/7) polypeptide
XX
XX XX SQ Sequence 2227 AA;
XX
XX Query Match 100.0%; Score 109; DB 5; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 VDDPRSEEDKRFESHIECRK 20
XX |||||
XX Db 779 VDDPRSEEDKRFESHIECRK 798
XX
XX RESULT 14
XX ABU08640
XX ID ABU08640 standard; protein; 2227 AA.
XX
XX AC ABU08640;
XX
XX XX 23-OCT-2003 (revised)
XX
XX DT 03-JUN-2003 (first entry)
XX
XX XX Attenuated (pass35) hepatitis A virus strain HM-175.
XX
XX DE Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
XX XX Hepatitis A virus; strain HM-175.
XX
XX XX US2002176869-A1.
XX
XX XX 28-NOV-2002.
XX
XX XX 29-APR-2002; 2002US-00135988.
XX
XX XX 18-SEP-1992; 92US-00947338.
XX

```

PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
PA (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93474.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Example 3; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;
XX
QY 1 VDDPRSEEDKRFESHIECRK 20
DB 779 VDDPRSEEDKRFESHIECRK 798
XX
RESULT 15
ABU08641
ID ABU08641 standard; protein; 2227 AA.
XX
AC ABU08641;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
DE Attenuated hepatitis A virus (4380) strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
OS Hepatitis A virus; strain HM-175.
XX
PN US2002176869-A1.
PN
PD 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
XX (EMER/) EMERSON S U.
XX (PURC/) PURCELL R H.
XX (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93475.
XX
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Disclosure; Page 45-51; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated human
CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;
XX
QY 1 VDDPRSEEDKRFESHIECRK 20
DB 779 VDDPRSEEDKRFESHIECRK 798
XX
RESULT 16
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX
AC ABU08639;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
DE Wild type human hepatitis A virus strain HM-175.
XX
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection.
XX
OS Hepatitis A virus; strain HM-175.
XX
PN US2002176869-A1.
PN
PD 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
XX
PR 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
XX (EMER/) EMERSON S U.
XX (PURC/) PURCELL R H.
XX (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
DR WPI; 2003-352605/02.
DR N-PSDB; ABX93473.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
PS Disclosure; Fig 6; 70pp; English.
XX
CC The invention describes a live hepatitis A virus (HAV) adapted to growth

CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)
XX
SQ Sequence 2227 AA;

Query Match 100.0%; Score 109; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-07; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
DB 779 VDDPRSEEDKRFESHIECKR 798

RESULT 17

AAP50230
ID AAP50230 standard; protein; 366 AA.

XX AC AAP50230;

XX DT 28-NOV-1991 (first entry)

XX DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).

XX KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;

XX KW diagnostic assay.

XX OS Hepatitis A virus.

XX PN EP138704-A.

XX PD 24-APR-1985.

XX PF 09-OCT-1984; 84EP-00402025.

XX PR 14-OCT-1983; 83US-00541836.

XX PR 02-MAR-1984; 84US-00585942.

XX PA (MERI) MERCK & CO INC.

XX PI Hughes JV, Scolnick EM, Tomassini JE;

XX DR WPI; 1985-100818/17.

XX DR N-PSDB; AAN50274.

XX PT New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus.

XX PS Claim 21; Page 46-48; 49pp; English.

XX CC VP1 is isolated by solubilisation of the intact virus in an aq. antonic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 daltons is sepd

XX SQ Sequence 366 AA;

Query Match 97.2%; Score 106; DB 1; Length 366;
Best Local Similarity 95.0%; Pred. No. 6.8e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
DB 288 VDDPRSEEDKRFESHIECKR 307

RESULT 18

AAP50116
ID AAP50116 standard; protein; 993 AA.

XX AC AAP50116;

XX XX

DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-SEP-1991 (first entry)
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
DE and VP-4.

XX KW Antigenic protein; immunogen; vaccine.

XX OS Hepatitis A virus; (strain CR326).

XX PN EP154587-A.

XX PD 11-SEP-1985.

XX PF 27-FEB-1985; 85EP-00400369.

XX PR 02-MAR-1984; 84US-00585818.

XX PA (MERI) MERCK & CO INC.

XX PI Linemeyer DL, Menke JG, Rueben RG, Mitra SW;

XX DR WPI; 1985-224964/37.

XX DR N-PSDB; AAN50139.

XX PT New nucleotide sequences coding for hepatitis A virus antigens - useful
PT for eliciting normal immune response and in vaccines for protecting
PT against the virus.

XX PS Example; Page 11-17; 32pp; English.

XX CC Within the sequence in AAN50139 is encoded the information necessary to
CC make the antigenic proteins of HAV. The sequences encoding for the
CC structural proteins begin at base 403. The key sub-unit sequences within
CC VP-1, designated Sequences I, II, III, IV, and V, start, respectively at
CC 1882, 1983, 1999, 2146, 2347. Other nucleotide sequences which are
CC valuable as encoding antigenic proteins are the sequences from base 1749
CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
CC vector for producing antigen protein. Sequences II-V are claimed. X in
CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)

XX SQ Sequence 993 AA;

Query Match 97.2%; Score 106; DB 1; Length 993;
Best Local Similarity 95.0%; Pred. No. 2e-07; Indels 0; Gaps 0;
Matches 19; Conservative 1; Mismatches 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||
DB 915 VDDPRSEEDKRFESHIECKR 934

RESULT 19

AAP50231
ID AAP50231 standard; protein; 993 AA.

XX AC AAP50231;

XX DT 28-NOV-1991 (first entry)

XX DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).

XX KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.

XX OS Hepatitis A virus.

XX FH Key Location/Qualifiers

FT Protein 628..993

FT /note= "claimed; X denotes translated stop codons and
FT unspecified triplets"
XX
XX EPI38704-A.
XX
XX 24-APR-1985.
XX PD
XX 09-OCT-1984; 84EP-00402025.
XX PF
XX 14-OCT-1983; 83US-00541836.
XX PR
XX 02-MAR-1984; 84US-00585942.
XX PR
XX (MERI) MERCK & CO INC.
XX PA
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
XX WPI; 1985-100818/17.
XX DR N-PSDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus.
XX
XX Disclosure; Page 17-23; 49pp; English.
XX
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 Daltons is sepd
XX
XX Sequence 993 AA;
SQ

Query Match 97.2%; Score 106; DB 1; Length 993;
Best Local Similarity 95.0%; Pred. No. 2e-07; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 915 VDDPRSEEDRRPFESHIECKR 934

RESULT 20
AAE19899
ID AAE19899 standard; protein; 2227 AA.
XX AC AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX DT
XX Hepatitis A virus (HAV) protein.
XX DE
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX OS
XX WO200213855-A2.
XX PN
XX 21-FEB-2002.
XX PD
XX 15-AUG-2001; 2001WO-IB001808.
XX PF
XX 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX XX
XX (TRIP-) TRIPEP AB.
XX PA
XX Sallberg M, Hultgren C;
XX PI
XX WPI; 2002-241837/29.
XX DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has

PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 82-87; 120pp; English.
XX PS
XX The invention relates to a composition comprising ribavirin and an
CC antigen preferably non structural 3 protein (NS3)/4A fragment of
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
CC sequence. The composition is useful for enhancing an immune response to a
CC hepatitis C antigen in humans, domestic, sport or pet species and as
CC vaccines for treating and preventing HCV infections. The composition is
CC also useful for treating viral, bacterial, fungal diseases and cancer.
CC The present sequence is hepatitis A virus (HAV) protein
XX
XX Sequence 2227 AA;
SQ

Query Match 97.2%; Score 106; DB 5; Length 2227;
Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDRRPFESHIECKR 798

RESULT 21
ABW00350
ID ABW00350 standard; protein; 2227 AA.
XX AC ABW00350;
XX
XX 15-JAN-2004 (first entry)
XX DT
XX Hepatitis A virus protein.
XX DE
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
KW virucide.
XX
XX Hepatitis A virus.
XX OS
XX US2002136740-A1.
XX PN
XX 26-SEP-2002.
XX PD
XX 15-AUG-2001; 2001US-00929955.
XX PF
XX 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR
XX (SALL/) SALLBERG M.
XX PA (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX PI
XX WPI; 2003-764978/72.
XX DR N-PSDB; AAD60867.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX
XX Claim 11; Page 45-51; Opp; English.
XX PS
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis A
CC virus protein
XX
XX Sequence 2227 AA;
SQ

Query Match 97.2%; Score 106; DB 7; Length 2227;

Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRPESHIECRK 20
|||||||:|||||
Db 779 VDDPRSEEDKRPESHIECRK 798

RESULT 22
AAP60066
ID AAP60066 standard; protein; 2227 AA.
XX AAP60066;
XX
XX 25-MAR-2003 (revised)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of viral L434 polypeptide encoded by the complete nucleotide
DE sequence of the HAV genome.
XX
XX Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
FH Key Location/Qualifiers
FT Region 1..245
FT /label= P1.1A
FT Region 246..491
FT /label= 1B
FT Region 492..836
FT /label= 1C
FT Region 837..980
FT /label= P2.2A
FT Region 981..1078
FT /label= 2B
FT Region 1079..1422
FT /label= 2C
FT Region 1423..1484
FT /label= P3.3A
FT Region 1485..1507
FT /label= 3B
FT Region 1508..1678
FT /label= 3C
FT Region 1679..2227
FT /label= 3D
XX
XX EP199480-A.
XX
XX 29-OCT-1986.
XX
XX 03-APR-1986; 86EP-00302465.
XX
XX 03-APR-1985; 85US-00719329.
XX
XX (CHIR) CHIRON CORP.
XX
XX Dina D, Potter SJ, Vannest GA, Caput D;
XX
XX WPI; 1986-286213/44.
XX
XX N-PSDB; AAN60080.
XX
XX Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.
XX
XX of vaccines and diagnostic probes.
XX
XX Claim 5; Fig 1; 18pp; English.
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
XX hepatitis A virus; transformed hosts may be used for expression of
XX polypeptides and fragments useful in vaccines without risk of infection
XX by the virus or in prodn. of particles which are capable of inducing
XX immunocompetent B cells for passive immunotherapy. Pref. epitope is
XX derived from AAs 445-657 or 792-848 of the HAV polypeptide sequence
XX (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 2227 AA;
Query Match 79.8%; Score 87; DB 1; Length 2227;
Best Local Similarity 80.0%; Pred. No. 0.0004;
Matches 16; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRPESHIECRK 20
|||||||:|||||
Db 779 VDDPRSEEDKRPESHIECRK 798

RESULT 23
AAR15629
ID AAR15629 standard; protein; 839 AA.
XX AAR15629;
XX
XX 24-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 17-MAR-1992 (first entry)
XX
DE Capsid region of cyno-HAV isolate CY-145.
XX
XX Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.
XX
OS Viruses.
XX
FH Key Location/Qualifiers
FT Protein 1..245
FT /label= VP0
FT Cleavage-site 245..246
FT Protein 246..491
FT /label= VP3
FT Active-site 315
FT Cleavage-site 491..492
FT Protein 492..791
FT /label= VP1
FT Active-site 593
FT Cleavage-site 791..792
FT Protein 792
FT /label= P2
FT /note= "incomplete"
XX
XX USN7678828-N.
XX
XX 12-NOV-1991.
XX
XX 03-APR-1991; 91US-00678828.
XX
XX 03-APR-1991; 91US-00678828.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Hainan OV, Margolis HS, Robertson BH, Brinton MH, Ebert JW;
XX
XX WPI; 1991-376737/51.
XX
XX N-PSDB; AAQ15180.
XX
XX Hepatitis A virus isolates and DNA - used to prepare vaccines for
XX preventing hepatitis A virus infection.
XX
XX Disclosure; Fig 3; 23pp; English.
XX
XX The sequence was deduced from the nucleotide sequence obt'd. by PCR
XX amplification of cyno-HAV viral RNA obt'd. from the stool of a cynomolgus
XX monkey with serologically and histologically confirmed spontaneous
XX hepatitis A. The sequence differs from the human HAV isolate HM175
XX (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci. USA 84, 2497-2501),
XX mainly in the VP3 and VP1 proteins. The Gln-Val pair at the VP3-VP1
XX cleavage site in the human isolate is replaced by a Gln-Thr pair in the

CC cyno-HAV. The other two cleavage sites are the same. Two residues have
 CC been identified as part of the immuno- dominant region (see feature
 CC table) and are different to those in the same position in human HAV. The
 CC protein and peptides derived from it can be used in the prepn. of
 CC vaccines for the prevention of HAV infection. See also AAR15056. (Note:
 CC Revised entry submitted to correct the patent number format of US
 CC Government-owned NTRIS applications to prevent clashes with ongoing US
 CC granted patent numbers. For further information please visit the Derwent
 CC web site at www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-
 CC MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
 CC field.) (Updated on 24-OCT-2003 to standardise OS field)
 CC
 XX SQ Sequence 839 AA;

Query Match 74.3%; Score 81; DB 2; Length 839;
 Best Local Similarity 82.4%; Pred. No. 0.0012;
 Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
 DB 779 VDDPTDEDRFESHIE 795

RESULT 24
 AAW42920
 ID AAW42920 standard; peptide; 20 AA.

XX AC AAW42920;

DT 28-APR-1998 (first entry)

XX Immunogenic Hepatitis A virus peptide YK-1313.

DE Immunogenic peptide; immunogenic epitope; VP1 protein; immune response;
 KW antibody.

XX Synthetic.

OS Hepatitis A virus.

XX WO9740147-A1.

XX 30-OCT-1997.

PF 18-APR-1997; 97WO-US006891.

PR 19-APR-1996; 96US-0015644P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

PI WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.

XX Claim 13; Page 111; 140pp; English.

XX Peptides AAW42906-21 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the VP1
 CC protein of HAV corresponding to amino acids 492-791. The present peptide
 CC is derived from amino acids 772-791, and has a reactivity of 37.5% with
 CC acute sera. Compositions containing the peptides can be used to induce an
 CC immune response to HAV in a mammal. The peptides can also be used to
 CC detect the presence of antibodies against HAV in mammalian serum. The
 CC peptides can also be used to make an antibody against HAV by
 CC administering the peptide to a mammal

XX SQ Sequence 20 AA;

Query Match 63.3%; Score 69; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.0014;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDKRF 13
 DB 8 VDDPRSEEDKRF 20

RESULT 25

AAB69437

ID AAB69437 standard; peptide; 21 AA.

XX AC AAB69437;

DT 20-APR-2001 (first entry)

XX Synthetic HAV VP1 peptide, SEQ ID NO: 37.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.

OS Synthetic.

XX WO200105824-A2.

XX 25-JAN-2001.

PF 14-JUL-2000; 2000WO-US019267.

PR 15-JUL-1999; 99US-0144412P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Fields HA, Khudyakov YE;

PI WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.

XX Claim 10; Page 91; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of antibodies against HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity

XX SQ Sequence 21 AA;

Query Match 63.3%; Score 69; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.0015;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRF 13

DB 8 VDDPRSEEDKRF 20

RESULT 26

AAP90612

ID AAP90612 standard; protein; 300 AA.
 AC AAP90612;
 DT 25-MAR-2003 (revised)
 DT 20-JUL-2000 (revised)
 DT 31-OCT-1989 (first entry)
 XX
 DE Portion of structural protein of VP1 of hepatitis A virus strain HM175.
 XX
 KW Hepatitis A virus; VP1; HM175 strain; DNA; immunogen; conjugate;
 KW vaccines; VP3.
 OS Hepatitis A virus.
 XX
 XX WO8906658-A.
 PN
 XX 27-JUL-1989.
 PD
 PF 13-JAN-1989; 89WO-US000097.
 XX
 XX 13-JAN-1989; 88US-00143395.
 PR
 XX (UYNC-) UNIV OF N CAROLINA.
 PA
 XX Lemon SM;
 PI
 XX WPI; 1989-233826/32.
 DR
 DR N-PSDB; AAN90399.
 XX
 XX Synthetic immunogenic peptide(s) corresp. to hepatitis A virus - derived
 PT from VP3 structural protein and used for vaccine prodn. and diagnosis.
 XX
 PS Disclosure; Fig 1; 29pp; English.
 XX
 XX Portion of VP1 (see AAN90399; depicting bases 2197-3106) structural
 CC protein of HM175 strain hepatitis A virus. The patent claims an
 CC immunogenic peptide corresp. to a region of VP3, and this can form an
 CC immunogenic conjugate with eg a portion of VP1 (contg. Ser-102). For VP3
 CC see AAN90400, AAP90613. (Revised entry issued to correct the sequence
 CC analysis breakdown.) (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 300 AA;
 Query Match 63.3%; Score 69; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VDDPRSEEDKQFE 13
 |||||
 Db 288 VDDPRSEEDKQFE 300
 RESULT 27
 AEM69971
 ID AEM69971 standard; protein; 409 AA.
 XX
 AC AEM69971;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 XX Photorhabdus luminescens protein sequence #3068.
 DE
 XX Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX
 OS Photorhabdus luminescens.
 XX
 XX WO200294867-A2.
 PN
 XX 28-NOV-2002.

XX 07-FEB-2002; 2002WO-IB003040.
 PF
 XX 07-FEB-2001; 2001FR-00001659.
 PR
 XX (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 XX Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 PT
 XX Claim 2; SEQ ID NO 3068; 1205pp; French.
 PS
 XX The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and
 CC antibacterials useful as insecticides, bactericides and fungicides. The
 CC genes, proteins, vectors containing the genes and Ab are also useful
 CC therapeutically (to treat microbial infection by bacteria or fungi that
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
 CC biopesticides. Other uses of the genes and the proteins are as virulence
 CC factors and for identifying targets of human diseases for which P.
 CC luminescens is a model (particularly plague and whooping cough). This
 CC sequence represents one of the isolated P. luminescens proteins
 XX
 SQ Sequence 409 AA;
 Query Match 45.9%; Score 50; DB 6; Length 409;
 Best Local Similarity 57.1%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 7 EEDKRFESHIECRK 20
 |. |||. :. |||
 Db 355 EDQKRTATVCECRK 368
 RESULT 28
 AAG57588
 ID AAG57588 standard; protein; 183 AA.
 XX
 AC AAG57588;
 XX
 DT 18-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 74230.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 XX EP1033405-A2.
 PN
 XX 06-SEP-2000.
 PD
 XX 25-FEB-2000; 2000EP-00301439.
 PF
 XX

PR	25-FEB-1999;	99US-0121825P.	PR	13-JUL-1999;	99US-0143542P.
PR	05-MAR-1999;	99US-0123180P.	PR	14-JUL-1999;	99US-0143624P.
PR	09-MAR-1999;	99US-0123548P.	PR	15-JUL-1999;	99US-0144005P.
PR	23-MAR-1999;	99US-0125788P.	PR	16-JUL-1999;	99US-0144085P.
PR	28-MAR-1999;	99US-0126264P.	PR	16-JUL-1999;	99US-0144086P.
PR	29-MAR-1999;	99US-0126785P.	PR	19-JUL-1999;	99US-0144332P.
PR	01-APR-1999;	99US-0127462P.	PR	19-JUL-1999;	99US-0144331P.
PR	06-APR-1999;	99US-0128234P.	PR	19-JUL-1999;	99US-0144332P.
PR	08-APR-1999;	99US-0128714P.	PR	19-JUL-1999;	99US-0144333P.
PR	16-APR-1999;	99US-0129845P.	PR	19-JUL-1999;	99US-0144334P.
PR	19-APR-1999;	99US-0130077P.	PR	19-JUL-1999;	99US-0144335P.
PR	23-APR-1999;	99US-0130449P.	PR	20-JUL-1999;	99US-0144352P.
PR	23-APR-1999;	99US-0130510P.	PR	20-JUL-1999;	99US-0144632P.
PR	28-APR-1999;	99US-0130891P.	PR	20-JUL-1999;	99US-0144884P.
PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	07-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145089P.
PR	11-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145112P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145115P.
PR	14-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134219P.	PR	23-JUL-1999;	99US-0145224P.
PR	18-MAY-1999;	99US-0134768P.	PR	26-JUL-1999;	99US-0145276P.
PR	19-MAY-1999;	99US-0134941P.	PR	27-JUL-1999;	99US-0145913P.
PR	20-MAY-1999;	99US-0135124P.	PR	27-JUL-1999;	99US-0145918P.
PR	21-MAY-1999;	99US-0135353P.	PR	27-JUL-1999;	99US-0145919P.
PR	24-MAY-1999;	99US-0135629P.	PR	27-JUL-1999;	99US-0145951P.
PR	25-MAY-1999;	99US-0136021P.	PR	28-JUL-1999;	99US-0145951P.
PR	27-MAY-1999;	99US-0136382P.	PR	02-AUG-1999;	99US-0146386P.
PR	28-MAY-1999;	99US-0136782P.	PR	02-AUG-1999;	99US-0146388P.
PR	01-JUN-1999;	99US-0137222P.	PR	02-AUG-1999;	99US-0146389P.
PR	03-JUN-1999;	99US-0137528P.	PR	02-AUG-1999;	99US-0146389P.
PR	04-JUN-1999;	99US-0137724P.	PR	03-AUG-1999;	99US-0147038P.
PR	07-JUN-1999;	99US-0137502P.	PR	04-AUG-1999;	99US-0147204P.
PR	08-JUN-1999;	99US-0138094P.	PR	04-AUG-1999;	99US-0147302P.
PR	10-JUN-1999;	99US-0138540P.	PR	05-AUG-1999;	99US-0147192P.
PR	10-JUN-1999;	99US-0138847P.	PR	05-AUG-1999;	99US-0147260P.
PR	14-JUN-1999;	99US-0139119P.	PR	06-AUG-1999;	99US-0147303P.
PR	16-JUN-1999;	99US-0139452P.	PR	06-AUG-1999;	99US-0147416P.
PR	16-JUN-1999;	99US-0139453P.	PR	09-AUG-1999;	99US-0147493P.
PR	17-JUN-1999;	99US-0139452P.	PR	09-AUG-1999;	99US-0147935P.
PR	18-JUN-1999;	99US-0139454P.	PR	09-AUG-1999;	99US-0148171P.
PR	18-JUN-1999;	99US-0139455P.	PR	10-AUG-1999;	99US-0148171P.
PR	18-JUN-1999;	99US-0139456P.	PR	11-AUG-1999;	99US-0148319P.
PR	18-JUN-1999;	99US-0139457P.	PR	11-AUG-1999;	99US-0148341P.
PR	18-JUN-1999;	99US-0139458P.	PR	12-AUG-1999;	99US-0148565P.
PR	18-JUN-1999;	99US-0139461P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139462P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139463P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US			

PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0158293P.
PR	13-OCT-1999;	99US-0158294P.
PR	13-OCT-1999;	99US-0158295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity 42.7%; Score 46.5; DB 3; Length 183;

Mismatches 10; Conservative 3; Indels 1; Gaps 1;

QY	1	VDDPRSEED-KRPESHIECR 19
Db	140	IDDSLDQETKRFVSHVEAR 159

RESULT 29

AAG57587

ID AAG57587 standard; protein; 187 AA.

AC AAG57587;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 74229.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

XX

PR 25-FEB-1999; 99US-0121825P.

PR

PR 05-MAR-1999; 99US-0123180P.

PR

PR 09-MAR-1999; 99US-0123548P.

PR

PR 23-MAR-1999; 99US-0125788P.

PR

PR 25-MAR-1999; 99US-0126264P.

PR

PR 29-MAR-1999; 99US-0126785P.

PR

PR 01-APR-1999; 99US-0127462P.

PR

PR

06-APR-1999;

99US-0128234P.

PR

08-APR-1999;

99US-0128714P.

PR

16-APR-1999;

99US-0129845P.

PR

19-APR-1999;

99US-0130077P.

PR

21-APR-1999;

99US-0130449P.

PR

23-APR-1999;

99US-0130510P.

PR

23-APR-1999;

99US-0130891P.

PR

28-APR-1999;

99US-0131449P.

PR

30-APR-1999;

99US-0132048P.

PR

30-APR-1999;

99US-0132407P.

PR

04-MAY-1999;

99US-0132484P.

PR

05-MAY-1999;

99US-0132485P.

PR

06-MAY-1999;

99US-0132486P.

PR

07-MAY-1999;

99US-0132487P.

PR

07-MAY-1999;

99US-0132863P.

PR

11-MAY-1999;

99US-0134256P.

PR

14-MAY-1999;

99US-0134218P.

PR

14-MAY-1999;

99US-0134219P.

PR

14-MAY-1999;

99US-0134221P.

PR

14-MAY-1999;

99US-0134370P.

PR

18-MAY-1999;

99US-0134768P.

PR

19-MAY-1999;

99US-0134941P.

PR

20-MAY-1999;

99US-0135124P.

PR

21-MAY-1999;

99US-0135353P.

PR

24-MAY-1999;

99US-0135629P.

PR

25-MAY-1999;

99US-0136021P.

PR

27-MAY-1999;

99US-0136392P.

PR

28-MAY-1999;

99US-0136782P.

PR

01-JUN-1999;

99US-0137222P.

PR

03-JUN-1999;

99US-0137528P.

PR

04-JUN-1999;

99US-0137502P.

PR

07-JUN-1999;

99US-0137724P.

PR

08-JUN-1999;

99US-0138094P.

PR

10-JUN-1999;

99US-0138540P.

PR

10-JUN-1999;

99US-0138847P.

PR

14-JUN-1999;

99US-0139119P.

PR

16-JUN-1999;

99US-0139452P.

PR

16-JUN-1999;

99US-0139453P.

PR

17-JUN-1999;

99US-0139492P.

PR

18-JUN-1999;

99US-0139454P.

PR

18-JUN-1999;

99US-0139455P.

PR

18-JUN-1999;

99US-0139456P.

PR

18-JUN-1999;

99US-0139457P.

PR

18-JUN-1999;

99US-0139458P.

PR

18-JUN-1999;

99US-0139459P.

PR

18-JUN-1999;

99US-0139460P.

PR

18-JUN-1999;

99US-0139461P.

PR

18-JUN-1999;

99US-0139462P.

PR

18-JUN-1999;

99US-0139463P.

PR

18-JUN-1999;

99US-0139750P.

PR

18-JUN-1999;

99US-0139763P.

PR

21-JUN-1999;

99US-0139817P.

PR

22-JUN-1999;

99US-0139899P.

PR

23-JUN-1999;

99US-0140353P.

PR

23-JUN-1999;

99US-0140354P.

PR

24-JUN-1999;

99US-0140695P.

PR

28-JUN-1999;

99US-0140823P.

PR

29-JUN-1999;

99US-0140931P.

PR

30-JUN-1999;

99US-0141287P.

PR

01-JUL-1999;

99US-0141842P.

PR

01-JUL-1999;

99US-0142154P.

PR

02-JUL-1999;

99US-0142055P.

PR

06-JUL-1999;

99US-0142390P.

PR

08-JUL-1999;

99US-0142803P.

PR

09-JUL-1999;

99US-0142920P.

PR

12-JUL-1999;

99US-0142977P.

PR

13-JUL-1999;

99US-0143542P.

PR

14-JUL-1999;

99US-0143624P.

PR

15-JUL-1999;

99US-0144005P.

PR

16-JUL-1999;

99US-0144085P.

PR

16-JUL-1999;

99US-0144086P.

PR

19-JUL-1999;

99US-0144325P.

PR

19-JUL-1999;

99US-0144331P.

PR

[illegible]

Accession	Score	Length	Indels	Mismatches	Conservative	Protein
PR 13-OCT-1999;	99US-0159294P.	42.7%;	DB 3;	Length 187;		
PR 13-OCT-1999;	99US-0159295P.	50.0%;	Pred. No. 46;			
PR 14-OCT-1999;	99US-0159329P.	10;	3;	Mismatches	6;	Indels 1; Gaps 1;
PR 14-OCT-1999;	99US-0159330P.					
PR 14-OCT-1999;	99US-0159331P.					
PR 14-OCT-1999;	99US-0159637P.					
PR 14-OCT-1999;	99US-0159638P.					
PR 14-OCT-1999;	99US-0159584P.					
PR 21-OCT-1999;	99US-0160741P.					
PR 21-OCT-1999;	99US-0160767P.					
PR 21-OCT-1999;	99US-0160768P.					
PR 21-OCT-1999;	99US-0160770P.					
PR 21-OCT-1999;	99US-0160814P.					
PR 22-OCT-1999;	99US-0160845P.					
PR 22-OCT-1999;	99US-0160980P.					
PR 22-OCT-1999;	99US-0160981P.					
PR 22-OCT-1999;	99US-0160989P.					
PR 25-OCT-1999;	99US-0161404P.					
PR 25-OCT-1999;	99US-0161405P.					
PR 25-OCT-1999;	99US-0161406P.					
PR 26-OCT-1999;	99US-0161359P.					
PR 26-OCT-1999;	99US-0161360P.					
PR 26-OCT-1999;	99US-0161361P.					
PR 28-OCT-1999;	99US-0161930P.					
PR 28-OCT-1999;	99US-0161932P.					
PR 28-OCT-1999;	99US-0161933P.					
PR 29-OCT-1999;	99US-0162142P.					
PR 13-OCT-1999;	99US-0144332P.					
PR 19-JUL-1999;	99US-0144333P.					
PR 19-JUL-1999;	99US-0144334P.					
PR 19-JUL-1999;	99US-0144335P.					
PR 20-JUL-1999;	99US-0144336P.					
PR 20-JUL-1999;	99US-0144337P.					
PR 20-JUL-1999;	99US-0144338P.					
PR 20-JUL-1999;	99US-0144339P.					
PR 20-JUL-1999;	99US-0144340P.					
PR 20-JUL-1999;	99US-0144341P.					
PR 20-JUL-1999;	99US-0144342P.					
PR 20-JUL-1999;	99US-0144343P.					
PR 20-JUL-1999;	99US-0144344P.					
PR 20-JUL-1999;	99US-0144345P.					
PR 20-JUL-1999;	99US-0144346P.					
PR 20-JUL-1999;	99US-0144347P.					
PR 20-JUL-1999;	99US-0144348P.					
PR 20-JUL-1999;	99US-0144349P.					
PR 20-JUL-1999;	99US-0144350P.					
PR 20-JUL-1999;	99US-0144351P.					
PR 20-JUL-1999;	99US-0144352P.					
PR 20-JUL-1999;	99US-0144353P.					
PR 20-JUL-1999;	99US-0144354P.					
PR 20-JUL-1999;	99US-0144355P.					
PR 20-JUL-1999;	99US-0144356P.					
PR 20-JUL-1999;	99US-0144357P.					
PR 20-JUL-1999;	99US-0144358P.					
PR 20-JUL-1999;	99US-0144359P.					
PR 20-JUL-1999;	99US-0144360P.					
PR 20-JUL-1999;	99US-0144361P.					
PR 20-JUL-1999;	99US-0144362P.					
PR 20-JUL-1999;	99US-0144363P.					
PR 20-JUL-1999						

PR	28-APR-1999;	99US-0131449P.	PR	21-JUL-1999;	99US-0144814P.
PR	30-APR-1999;	99US-0132048P.	PR	21-JUL-1999;	99US-0145086P.
PR	30-APR-1999;	99US-0132407P.	PR	21-JUL-1999;	99US-0145088P.
PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	06-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	07-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134258P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135629P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	03-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	04-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147202P.
PR	01-JUN-1999;	99US-0137222P.	PR	05-AUG-1999;	99US-0147192P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147260P.
PR	04-JUN-1999;	99US-0137502P.	PR	06-AUG-1999;	99US-0147303P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147416P.
PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148584P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139458P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	18-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149829P.
PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	22-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	01-JUL-1999;	99US-0141842P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0142154P.	PR	16-SEP-1999;	99US-0154039P.
PR	02-JUL-1999;	99US-0142055P.	PR	20-SEP-1999;	99US-0154779P.
PR	06-JUL-1999;	99US-0142390P.	PR	22-SEP-1999;	99US-0155139P.
PR	08-JUL-1999;	99US-0142803P.	PR	23-SEP-1999;	99US-0155486P.
PR	09-JUL-1999;	99US-0142920P.	PR	24-SEP-1999;	99US-0155659P.
PR	12-JUL-1999;	99US-0142977P.	PR	28-SEP-1999;	99US-0156458P.
PR	13-JUL-1999;	99US-0143542P.	PR	29-SEP-1999;	99US-0156596P.
PR	14-JUL-1999;	99US-0143624P.	PR	04-OCT-1999;	99US-0157117P.
PR	15-JUL-1999;	99US-0144005P.	PR	05-OCT-1999;	99US-0157753P.
PR	16-JUL-1999;	99US-0144085P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158312P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144684P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144684P.	PR	14-OCT-1999;	99US-0159638P.

```
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 42.7%; Score 46.5; DB 3; Length 200;
Best Local Similarity 50.0%; Pred. No. 49;
Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 1 VDDPRSEED-KREESHIECR 19
Db 157 IDDSLQETKRFVSHVEAR 176

Search completed: May 11, 2004, 13:37:21
Job time : 52 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:30:01 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-09-171-432a-38

Perfect score: 109

Sequence: 1 VDDPRSEDKRFESHCCK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR 78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	56	PQ0431	genome polypotein
2	109	100.0	319	JH0135	genome polypotein
3	109	100.0	1358	A03905	genome polypotein
4	109	100.0	2227	1 GNNYHM	genome polypotein
5	109	100.0	2227	1 GNNYMK	genome polypotein
6	109	100.0	2227	1 GNNYHB	genome polypotein
7	106	97.2	56	PQ0427	genome polypotein
8	106	97.2	56	PQ0429	genome polypotein
9	106	97.2	341	S04137	genome polypotein
10	106	97.2	852	1 GNNYHA	genome polypotein
11	106	97.2	2227	1 GNNYHR	genome polypotein
12	99	90.8	56	PQ0430	genome polypotein
13	96	88.1	56	PQ0432	genome polypotein
14	95	87.2	56	PQ0428	genome polypotein
15	91	83.5	56	2 PQ0434	genome polypotein
16	84	77.1	56	2 PQ0437	genome polypotein
17	84	77.1	56	2 PQ0436	genome polypotein
18	84	77.1	2230	1 GNNYSA	genome polypotein
19	81	74.3	55	2 PQ0433	genome polypotein
20	81	74.3	55	2 PQ0435	genome polypotein
21	81	74.3	839	1 GNNYS2	genome polypotein
22	60	55.0	300	2 A33327	genome polypotein
23	51	46.8	1548	2 T25808	hypothetical prote
24	50.5	46.3	754	2 T25851	hypothetical prote
25	50	45.9	1663	1 C3RT	complement C3 prec
26	50	45.9	1666	1 C3GP	complement C3 - Af
27	48	44.0	322	2 A32329	legumin B - fava b
28	48	44.0	564	2 S37241	hypothetical prote
29	46.5	42.7	454	2 B84699	hypothetical prote

30	46	42.2	242	2	E83077	conserved hypothet
31	46	42.2	308	2	T17524	hypothetical prote
32	46	42.2	1663	1	C3MS	complement C3 prec
33	45.5	41.7	650	2	T36419	hypothetical prote
34	45	41.3	209	2	T20155	hypothetical prote
35	45	41.3	480	2	T20154	hypothetical prote
36	45	41.3	586	1	E69314	replication licens
37	45	41.3	741	2	S73827	hypothetical prote
38	44	40.4	222	2	E71507	probable L4 riboso
39	44	40.4	234	2	B83065	conserved hypothet
40	44	40.4	433	2	A23850	phosphoryruvate hy
41	44	40.4	490	2	C84091	hypothetical prote
42	44	40.4	499	2	T04730	cytochrome P450 ho
43	44	40.4	593	2	A81021	para-aminobenzoate
44	44	40.4	598	2	D81965	probable para-amin
45	44	40.4	782	2	G96698	hypothetical prote
46	44	40.4	957	2	T50789	hypothetical prote
47	43.5	39.9	354	2	T12741	glucose dehydrogen
48	43.5	39.9	801	1	S00943	glutamate CoA-tr
49	43	39.4	266	2	S51052	hypothetical prote
50	43	39.4	281	2	H71308	hypothetical prote
51	43	39.4	318	2	A84466	hypothetical prote
52	43	39.4	352	2	G69774	hypothetical prote
53	43	39.4	478	2	S33886	DNA-directed RNA p
54	43	39.4	497	2	T04731	cytochrome P450 ho
55	43	39.4	500	2	C85441	cytochrome P450-11
56	43	39.4	500	2	T52174	aconitate hydratase
57	43	39.4	780	2	T52543	hypothetical prote
58	43	39.4	963	2	T19140	spectrin beta chain
59	43	39.4	2137	1	SJHUB	probable pre-tRNA
60	42.5	39.0	978	2	T40803	hypothetical prote
61	42	38.5	176	2	A86441	assemblin, striate
62	42	38.5	276	2	T08178	hypothetical 45.1
63	42	38.5	408	2	H65137	omega-3 fatty acid
64	42	38.5	438	2	T15039	CAMP-dependent pro
65	42	38.5	464	2	A32461	complement C9 prec
66	42	38.5	559	1	C9HU	2-aminobenzoate-Co
67	42	38.5	603	2	S22402	hypothetical prote
68	42	38.5	638	2	T13691	hypothetical prote
69	42	38.5	679	2	S48437	toxin secretion intia
70	42	38.5	704	2	H82381	translation initiat
71	42	38.5	918	2	T38786	polyketide synthase
72	41.5	38.1	230	2	H70349	translation elonga
73	41	37.6	113	2	JC5857	anaerobic glycerol
74	41	37.6	396	1	A54536	probable membrane
75	41	37.6	415	2	D86937	hypothetical prote
76	41	37.6	432	2	D86937	probable cytochrom
77	41	37.6	434	2	S37907	probable proline o
78	41	37.6	481	2	B96691	probable FAD bindi
79	41	37.6	492	2	T41543	hypothetical prote
80	41	37.6	571	2	T39378	hypothetical prote
81	41	37.6	607	2	T20796	hypothetical prote
82	41	37.6	614	2	T20795	peptidase [impor
83	41	37.6	654	2	AD3183	hypothetical prote
84	41	37.6	691	2	S48330	hypothetical prote
85	41	37.6	709	2	T00664	hypothetical prote
86	41	37.6	1021	2	G01202	NaCl electroneutra
87	41	37.6	1780	2	T17272	hypothetical prote
88	41	37.6	1805	2	A34736	nestin - rat
89	41	37.6	1882	1	GNVTR	genome polypotein
90	41	37.6	6642	2	T29757	protein UNC-89 C
91	40.5	37.2	359	2	T33115	hypothetical prote
92	40.5	37.2	470	2	D90895	probable aldehyde
93	40.5	37.2	470	2	D85722	probable aldehyde
94	40	36.7	138	2	T43419	ribosomal protein
95	40	36.7	140	2	T39778	40s ribosomal prot
96	40	36.7	198	2	D88098	protein f06b4.5 [i
97	40	36.7	198	2	T32025	hypothetical prote
98	40	36.7	205	2	JC7975	Hiv accessory prot
99	40	36.7	219	2	AD2776	transcription regu
100	40	36.7	227	2	S44836	K2D10.4 protein -

A;Molecule type: genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: ENBL:M16632; NID:G329594; PIDN:AAA4571.1; PID:G329595
A;Note: Submitted to GenBank, August 1987
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F;1-245/Product: coat protein 1A #status predicted <P1A>
F;246-491/Product: coat protein 1B #status predicted <P1B>
F;492-836/Product: coat protein 1C #status predicted <P1C>
F;837-980/Product: core protein 2A #status predicted <P2A>
F;981-1076/Product: core protein 2B #status predicted <P2B>
F;1077-1422/Product: core protein 2C #status predicted <P2C>
F;1423-1484/Product: core protein 3A #status predicted <P3A>
F;1485-1507/Product: core protein 3B #status predicted <P3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||:|||||
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 6
GNVHB
genome polyprotein - human hepatitis A virus (strain MBB)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C;Accession: J50303
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, Virus Res. 8, 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat A;Reference number: J50303; MUID:88045071; PMID:2823500
A;Accession: J50303
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: ENBL:M20273
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: core protein 2A #status predicted <P2A>
F;981-1108/Product: core protein 2B #status predicted <P2B>
F;1109-1438/Product: core protein 2C #status predicted <P2C>
F;1439-1496/Product: protein 3A #status predicted <P3A>
F;1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F;1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||:|||||
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 7
PQ0427
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0427
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel,

J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0427
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
C;Comment: This protein is from the VP1/2A Junction region.
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 97.2%; Score 106; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.5e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

RESULT 8
PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0429
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0429
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
C;Note: This protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 97.2%; Score 106; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.5e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||:|||||
Db 16 VDDPRSEEDKRFESHIECKR 35

RESULT 9
S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C;Species: human hepatitis A virus
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C;Accession: S04137
R;Andonov, A.P.; Lau, P.C.K.; Chaudhary, R. Nucleic Acids Res. 17, 3594, 1989
A;Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A;Reference number: S04137; MUID:89263805; PMID:2542903
A;Accession: S04137
A;Molecule type: mRNA
A;Residues: 1-341 <AND>
A;Cross-references: ENBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576
C;Genetics: C;Gene: VP1
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; polyprotein
F;2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 97.2%; Score 106; DB 2; Length 341;
Best Local Similarity 95.0%; Pred. No. 1e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
|||||:|||||:|||||
Db 283 VDDPRSEEDKRFESHIECKR 302

PQ0430
genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 90.8%; Score 99; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 1.8e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 16 VDDPRSEEDRRFESHIEGK 35
RESULT 13
PQ0432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0432
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 88.1%; Score 96; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 5.2e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 16 VDDPRSEEDRRFESHIESRK 35
RESULT 14
PQ0428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0428
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 87.2%; Score 95; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 7.5e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20

RESULT 10
GNNVHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A>Note: host Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03904
R:Linemeyer, D.L.; Wenke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793
A:Accession: A03904
A:Molecule type: genomic RNA
A:Residues: 1-852 <LIN>
A:Cross-references: EMBL:M10033; NID:G329592; PIDN:AAA45470.1; PID:G329593
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>
Query Match 97.2%; Score 106; DB 1; Length 852;
Best Local Similarity 95.0%; Pred. No. 2.6e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDRRFESHIECKR 798
RESULT 11
GNNVHR
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A>Note: host Homo sapiens (man)
C>Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nee
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:Title: Primary structure and gene organization of human hepatitis A virus.
A:Reference number: A03903; MUID:85190549; PMID:2986127
A:Accession: A03903
A:Molecule type: genomic RNA
A:Residues: 1-2227 <NAJ>
A:Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-980/Product: core protein 2A #status predicted <C2A>
F:981-1076/Product: core protein 2B #status predicted <C2B>
F:1077-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1484/Product: protein 3A #status predicted <C3A>
F:1485-1507/Product: protein 3B #status predicted <C3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>
Query Match 97.2%; Score 106; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 7.3e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDRRFESHIECKR 798
RESULT 12

```

Query Match      77.1%; Score 84; DB 2; Length 56;
Best Local Similarity 94.1%; Pred. No. 3.8e-06;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db
1 VDDPRSEEDKRFESHIE 17
16 VDDPRSAEDKRFESHIE 32

RESULT 18
GNNYSA
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein
C:Species: simian hepatitis A virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JIPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: Genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:G222597; PIDN:BA00766.1; PID:G222598
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and
A:Reference number: JQ1080; MUID:91311420; PMID:1649901
A:Contents: annotation
A:Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: Genomic RNA
A:Residues: 1750-2164 <BAL2>
A:Cross-references: EMBL:X15461; NID:G61971; PIDN:CAA33490.1; PID:G930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
FEBS Lett. 247, 425-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: Genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match      77.1%; Score 84; DB 1; Length 2230;
Best Local Similarity 94.1%; Pred. No. 0.00019;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db
1 VDDPRSEEDKRFESHIE 17
783 VDDPRSAEDKRFESHIE 799

RESULT 19
PQ0433
genome polyprotein - human hepatitis A virus (strain PA21) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0433

```

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0433
A;Molecule type: mRNA
A;Residues: 1-55 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 74.3%; Score 81; DB 2; Length 55;
Best Local Similarity 82.4%; Pred. No. 1.1e-05;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
|||||:|||||
DB 15 VDDPRTDEDRKFESHIE 31

RESULT 20
PQ0435
genome polyprotein - human hepatitis A virus (strain Cy145) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0435
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0435
A;Molecule type: mRNA
A;Residues: 1-55 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 74.3%; Score 81; DB 2; Length 55;
Best Local Similarity 82.4%; Pred. No. 1.1e-05;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
|||||:|||||
DB 16 VDDPRTDEDRKFESHIE 32

RESULT 21
GNYS2
genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
N;Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein
C;Species: simian hepatitis A virus
A;Note: host Macaca fascicularis (cynomolgus macaque)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C;Accession: JQ1180
R;Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A. J. Gen. Virol. 72, 1685-1689, 1991
A;Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus mac
A;Reference number: JQ1180; MUID:91311421; PMID:1649902
A;Accession: JQ1180
A;Molecule type: genomic RNA
A;Residues: 1-839 <NAI>
A;Cross-references: GB:M59286; NID:G329599; PIDN:AAA45473.1; PID:G555083
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; glycoprotein; polyprotein
F;1-23/Product: coat protein 1A #status predicted <VP0>
F;24-245/Product: coat protein 1B #status predicted <VP3>
F;246-491/Product: coat protein 1C #status predicted <VP1>
F;492-839/Product: coat protein 2A (fragment) #status predicted <P2P>
F;461,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.3%; Score 81; DB 1; Length 839;
Best Local Similarity 82.4%; Pred. No. 0.00019;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
|||||:|||||
DB 779 VDDPRTDEDRKFESHIE 795

RESULT 22
A33327
genome polyprotein - simian hepatitis A virus (strain PA21) (fragment)
C;Species: simian hepatitis A virus
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 15-Nov-1996
C;Accession: A33327
R;Brown, E.A.; Jansen, R.W.; Lemon, S.M. J. Virol. 63, 4932-4937, 1989
A;Title: Characterization of a simian hepatitis A virus (HAV): antigenic and genetic com
A;Reference number: A33327; MUID:90012354; PMID:2552172
A;Accession: A33327
A;Molecule type: genomic RNA
A;Residues: 1-300 <BRO>
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; polyprotein
F;1-300/Product: coat protein 1D #status predicted <CPD>

Query Match 55.0%; Score 60; DB 2; Length 300;
Best Local Similarity 76.9%; Pred. No. 0.12;
Matches 10; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 13
|||||:|||||
DB 288 VDDPRTDEDRKFESHIE 300

RESULT 23
T25808
hypothetical protein K08B12.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
C;Accession: T25808
R;Becker, M.; Wohldmann, P. submitted to the EMBL Data Library, April 1997
A;Description: The sequence of C. elegans cosmid K08B12.
A;Reference number: Z20091
A;Accession: T25808
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1548 <BEC>
A;Cross-references: EMBL:U97001; PIDN:AAB52260.1; GSPDB:GN00023; CESP:K08B12.5
A;Experimental source: strain Bristol N2; clone K08B12
C;Genetics:
A;Gene: CESP:K08B12.5
A;Map position: 5
A;Introns: 39/1; 66/1; 156/3; 278/2; 391/3; 628/1; 1022/2; 1273/3; 1326/1; 1423/2; 1489/3
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
F;956-1005/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 46.8%; Score 51; DB 2; Length 1548;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIE 16
|||||:|||||
DB 479 ENPRSEEDRYESTI 493

RESULT 24
T25551
hypothetical protein C17H11.2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
C;Accession: T25551
R;Johnson, D. submitted to the EMBL Data Library, December 1996
A;Description: The sequence of C. elegans cosmid C17H11.

A;Reference number: Z20049
A;Accession: T25551
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-754 <JOH>
A;Cross-references: EMBL:U80847; PIDN:AAB37983.1; GSPDB:GN00028; CESP:C17H11.2
A;Experimental source: strain Bristol N2; clone C17H11
C;Genetics:
A;Gene: CESP:C17H11.2
A;Map position: X
A;Introns: 155/1; 246/3; 288/3; 368/3; 631/1
C;Superfamily: Caenorhabditis elegans hypothetical protein C17H11.2

Query Match 46.3%; Score 50.5; DB 2; Length 754;
Best Local Similarity 37.5%; Pred. No. 9.2;
Matches 9; Conservative 7; Mismatches 3; Indels 5; Gaps 1;

QY 2 DDPSEEDK----RFESHIECR 20
DB 580 DEPLTEQRIHKRFYEVKCR 603

RESULT 25
CPT
complement C3 precursor - rat
N;Alternate names: 37K phospholipase A2 inhibitory protein
N;Contains: alternative-complement-pathway C3/C5 convertase (BC 3.4.21.47) C3b subunit;
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Dec-1992 #sequence revision 07-Oct-1994 #text change 18-Jun-1999
C;Accession: S15764; A54562; A01260; B35979; A35579; PNO566; A32281; S08692
R;Mitsumi, Y.; Kohda, M.; Ikehara, Y.
Nucleic Acids Res. 18, 2178, 1990
A;Title: Nucleotide and deduced amino acid sequence of rat complement C3.
A;Reference number: S15764; MUID:90245672; PMID:2336397
A;Accession: S15764
A;Molecule type: mRNA
A;Residues: 1-1663 <MIS>
A;Cross-references: EMBL:X52477; NID:g56953; PIDN:CAA36716.1; PID:g56954
R;Sundstrom, S.A.; Komm, B.S.; Ponce-de-Leon, H.; Yi, Z.; Teuscher, C.; Lyttle, C.R.
J. Biol. Chem. 264, 16941-16947, 1999
A;Title: Estrogen regulation of tissue-specific expression of complement C3.
A;Reference number: A54562; MUID:189380332; PMID:2674144
A;Accession: A54562
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: P',1316-1595 <SUN>
A;Cross-references: GB:M29866; NID:Q203200; PIDN:AAA40837.1; PID:g554423
R;Jacobs, J.W.; Rubin, J.S.; Huggli, T.E.; Bogardt, R.A.; Mariz, I.K.; Daniels, J.S.; Dau
Biochemistry 17, 5031-5038, 1978
A;Title: Purification, characterization, and amino acid sequence of rat anaphylatoxin (C
A;Reference number: A01260; MUID:79062262; PMID:309768
A;Accession: A01260
A;Molecule type: protein
A;Residues: 671-703, 'K', 705-720, 'KL', 723-748 <JAC>
A;Note: three disulfide bonds are present
R;Suwa, Y.; Kudo, I.; Imazumi, A.; Okada, M.; Kamimura, T.; Suzuki, Y.; Chang, H.W.; Ha
Proc. Natl. Acad. Sci. U.S.A. 87, 2395-2399, 1990
A;Title: Proteinaceous inhibitors of phospholipase A-2 purified from inflammatory sites
A;Reference number: A35979; MUID:90207203; PMID:2320562
A;Accession: B35979
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'X', 998-1005 <SUW>
A;Accession: A35979
A;Molecule type: protein
A;Residues: 'X', 961-962, 'P', 964-969 <SUZ>
R;Kakagawa, H.; Komorita, N.
Biochem. Biophys. Res. Commun. 194, 1181-1187, 1993
A;Title: Complement component C3-derived neutrophil chemotactic factors purified from ex
A;Reference number: PNO566; MUID:93356786; PMID:8352775
A;Accession: PNO567
A;Molecule type: protein
A;Residues: 568-592 <NAK>

A;Note: amino end of a C3-derived peptide designated exudate neutrophil chemotactic factor
A;Accession: PNO566
A;Molecule type: protein
A;Residues: 671-687 <NA2>
A;Note: amino end of peptide designated neutrophil chemotactic factor 1 and probably ide
R;Kuivane, P.C.; Capulong, R.B.; Harkins, R.N.; DeSombre, E.R.
Biochem. Biophys. Res. Commun. 158, 898-905, 1989
A;Title: The estrogen-responsive 110K and 74K rat uterine secretory proteins are structu
A;Reference number: A32281; MUID:89149812; PMID:2645873
A;Accession: A32281
A;Molecule type: protein
A;Residues: 25-41 <KUI>
A;Experimental source: 17beta-estradiol-stimulated uterus of immature rat
A;Note: the authors treat this 74K uterine secretory protein, identical as far as sequen
ent
C;Comment: Complement C3 contains two chains, formed by removal of four residues and lin
alternative complement pathways, releases the C3a anaphylatoxin from the amino end of t
native-complement-pathway C3/C5 convertase.
C;Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C;Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign pa
e classical-complement-pathway C3/C5 convertase. The activity of C3b is regulated by pro
C;Comment: The major site of synthesis of this plasma protein is the liver.
C;Superfamily: alpha-2-macroglobulin
C;Keywords: acute phase; chemotaxis; complement alternate pathway; complement pathway; g
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-666/Product: complement C3 and C3b beta chain #status predicted <C3BB>
F;25-666,671-1663/Product: complement C3 #status predicted <CC3>
F;25-666,749-1663/Product: complement C3b #status predicted <C3B>
F;671-1663/Product: complement C3 alpha chain #status predicted <CC3A>
F;671-748/Product: C3a anaphylatoxin #status experimental <C3T>
F;749-1663/Product: complement C3b alpha' chain #status predicted <C3BA>
F;946-1303/Product: C3dk fragment #status predicted <CDK>
F;1002-1303/Product: C3d fragment #status predicted <C3D>
F;1424-1457/Region: properdin binding
F;558-616,626-661,693-720,694-727,707-728,873-1513,1101-1158,1358-1489,1389-1458,1506-151
F;748-749/Cleavage site: Arg-Ser (C3 convertase) #status predicted
F;939,1617/Binding site: carbonyl (Asn) (covalent) #status predicted
F;1010-1013/Cross-link: thioester (Cys-Gln) #status predicted
F;1303-1304/Cleavage site: Arg-Ser (complement factor I) #status predicted
F;1320-1321/Cleavage site: Arg-Ser (complement factor I) #status predicted

Query Match 45.9%; Score 50; DB 1; Length 1663;
Best Local Similarity 38.9%; Pred. No. 25;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIECR 19
DB 1574 DEVQAGQERFISHVKCR 1591

RESULT 26
C3GP
complement C3 precursor - guinea pig
N;Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) C3b subunit; (C
C;Species: Cavia porcellus (guinea pig)
C;Date: 07-Feb-1992 #sequence revision 07-Oct-1994 #text change 18-Jun-1999
C;Accession: A37156; S03375; A20342; D20342; D20342; A31222
R;Auerbach, H.S.; Burger, R.; Dodds, A.; Colten, H.R.
J. Clin. Invest. 86, 96-106, 1990
A;Title: Molecular basis of complement C3 deficiency in guinea pigs.
A;Reference number: A37156; MUID:90307998; PMID:19731176
A;Accession: A37156
A;Molecule type: mRNA
A;Residues: 1-1666 <AUE>
A;Cross-references: GB:M34054; NID:g191262; PIDN:AAA37038.1; PID:g305335
R;Gerard, N.P.; Lively, M.O.; Gerard, C.
Protein Seq. Data Anal. 1, 473-478, 1988
A;Title: Amino acid sequence of guinea pig C3a anaphylatoxin.
A;Reference number: S03375; MUID:89113342; PMID:3064079
A;Accession: S03375
A;Molecule type: protein
A;Residues: 676-730, 'N', 732-752 <GER>
A;Experimental source: complement-activated guinea pig serum

A:Note: form isolated is inactive C3a anaphylatoxin and is missing the carboxyl-terminal R;Thomas, M.L.; Tack, B.F.
Biochemistry 22, 942-947, 1983
A:Title: Identification and alignment of a thiol ester site in the third component of guinea pig C3.
A:Reference number: A90479, MUID:83178889; PMID:6838833
A:Accession: A20342
A:Molecule type: protein
A:Residues: 676-687 <TH>
A:Accession: D20342
A:Molecule type: protein
A:Residues: 993-1012,1014-1017,'E',1019-1030,'Y' <TH2>
R;Goldberger, G.; Thomas, M.L.; Tack, B.F.; Williams, J.; Colten, H.R.; Abraham, G.N.
J. Biol. Chem. 256, 12617-12619, 1981
A:Title: NH2-terminal structure and cleavage of guinea pig pro-C3, the precursor of the C3a anaphylatoxin.
A:Reference number: A20342; MUID:82075767; PMID:6458605
A:Accession: C20342
A:Molecule type: protein
A:Residues: 23-38 <GOL>
C:Comment: Complement C3 contains two chains, formed by removal of four residues and linkage of the two chains by a disulfide bond. The C3a anaphylatoxin is released from the amino end of the C3 molecule.
C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.
C:Comment: C3b, with its highly reactive thiol group, binds to the surface of foreign particles and is cleaved to C3a and C3b. The activity of C3b is regulated by protein C.
C:Comment: The major site of synthesis of this plasma protein is the liver.
C:Superfamily: alpha-2-macroglobulin
C:Keywords: acute phase; complement alternate pathway; complement pathway; glycoprotein; F1-22/Domain; signal sequence #status predicted <SIG>
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-671/Product: Complement C3 and C3b beta chain #status predicted <C3BB>
F:23-671, 676-1666/Product: complement C3 #status predicted <C3>
F:23-671, 754-1666/Product: complement C3b #status predicted <C3B>
F:1676-1666/Product: complement C3 alpha chain #status predicted <C3A>
F:1676-753/Product: C3a anaphylatoxin #status predicted <C3T>
F:754-1666/Product: complement C3b alpha' chain #status predicted <C3BA>
F:951-1308/Product: C3d fragment #status predicted <CDK>
F:1007-1308/Product: C3d fragment #status predicted <C3D>
F:1429-1461/Region: properdin binding
F:7557-821,630-666,698-725,699-732,712-733,878-1517,1106-1163,1363-1493,1394-1462,1510-1517
F:7557-821,630-666,698-725,699-732,712-733,878-1517,1106-1163,1363-1493,1394-1462,1510-1517
F:7557-821,630-666,698-725,699-732,712-733,878-1517,1106-1163,1363-1493,1394-1462,1510-1517
F:944-1620/Binding site: Arg-Ser (C3 convertase) #status predicted
F:1015-1018/Cross-link: carboxylate (Asn) (covalent) #status predicted
F:1015-1018/Cross-link: thioester (Cys-Gln) #status experimental
F:1308-1309/Cleavage site: Arg-Ser (complement factor I) #status predicted
F:1325-1326/Cleavage site: Arg-Ser (complement factor I) #status predicted
Query Match 45.9%; Score 50; DS 1; Length 1666;
Best Local Similarity 44.4%; Pred. No. 25;
Matches 8; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
QY 2 DDPREEDKRFESHIECR 19
DB 1577 DEVQAGKRRFTHSHIKR 1594

RESULT 27
A32329
complement C3 - African clawed frog (fragment)
C1:Species: Xenopus laevis (African clawed frog)
C2:Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 07-Oct-1994
C3:Accession: A32329
R:Grossberger, D.; Marcuz, A.; Du Pasquier, L.; Lambris, J.D.
Proc. Natl. Acad. Sci. U.S.A. 86, 1323-1327, 1989
A:Title: Conservation of structural and functional domains in complement component C3 of
A:Reference number: A32329; M01D:89145234; PMID:2919181
A:Accession: A32329
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-322 <GRO>
A:Cross-references: GB:M21790; GB:J04493
C:Comment: Complement C3 contains two chains, formed by removal of four residues and lin-
nate-complement pathway C3/C5 convertase, releases the C3a anaphylatoxin from the amino end of t
rnative-complement-pathway C3/C5 convertase.
C:Comment: C3a anaphylatoxin is a vasoactive peptide and a mediator of inflammation.


```
RESULT 30
E83077
conserved hypothetical protein PA4543 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83077
R:Stover, C.K.; Pham, X.Q.; Brwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: E83077
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <STO>
A:Cross-references: GB:AE004868; GB:AE004091; NID:g9950785; PIDN:AAG07931.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4543
C:Superfamily: Streptomyces coelicolor hypothetical protein SC4A10.14c

Query Match 42.2%; Score 46; DB 2; Length 242;
Best Local Similarity 45.0%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 DDPRSEED--KRFESHIECR 19
Db 42 DDPRAVEENRRLLTELECR 61

Search completed: May 11, 2004, 13:41:14
Job time : 12.75 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:24:26 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDFRSEEDKRFESHIECRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	109	100.0	2226	1	POLG_HPAV2
2	109	100.0	2226	1	POLG_HPAV4
3	109	100.0	2226	1	POLG_HPAV8
4	109	100.0	2227	1	POLG_HPAVH
5	109	100.0	2227	1	POLG_HPAVM
6	106	97.2	341	1	POLG_HPAV1
7	106	97.2	852	1	POLG_HPAVC
8	106	97.2	2227	1	POLG_HPAVL
9	88	80.7	808	1	POLG_HPAVG
10	84	77.1	2230	1	POLG_HPAVS
11	81	74.3	839	1	POLG_HPAVT
12	50	45.9	1663	1	CO3_RAT
13	50	45.9	1666	1	CO3_CAVPO
14	48	44.0	323	1	CO3_XENLA
15	46	42.2	242	1	Y983_PSEAE
16	46	42.2	1663	1	CO3_MOUSE
17	45	41.3	741	1	YD35_MYCPN
18	44	40.4	222	1	RL4_CHLTR
19	44	40.4	276	1	TRME_HUMAN
20	44	40.4	1101	1	FURC_DROME
21	43.5	39.9	801	1	DHGA_ACICA
22	43	39.4	265	1	GCTB_ACIFE
23	43	39.4	267	1	SURE_METMA
24	43	39.4	478	1	RPB1_EUPOC
25	43	39.4	502	1	C8LD_ARATH
26	43	39.4	507	1	DNAL_FUSNN
27	43	39.4	780	1	ACON_HUMAN
28	43	39.4	963	1	YQ36_CABEL
29	43	39.4	2137	1	SPCB_HUMAN
30	42	38.5	269	1	SPAS_CHLEU
31	42	38.5	276	1	SPAS_CHLEU
32	42	38.5	408	1	RTCB_ECOLI
33	42	38.5	464	1	AK15_RAT

34	38.5	559	1	CO9_HUMAN	P02748	homo sapien
35	38.5	639	1	DNAB_SHEON	Q8ht7	shewanella
36	38.5	679	1	YIS3_YEAST	P40563	saccharomyc
37	38.5	918	1	IF38_SCHPO	O14164	schizosacch
38	37.6	267	1	SURE_METAC	Q8tug3	methanosarc
39	37.6	396	1	EFTU_FLESI	P26184	flexistipes
40	37.6	434	1	YK12_YEAST	P36080	saccharomyc
41	37.6	492	1	PROD_SCHPO	O74524	schizosacch
42	37.6	571	1	YB63_SCHPO	Q09744	schizosacch
43	37.6	691	1	YI04_YEAST	P40460	saccharomyc
44	37.6	925	1	NRP2_RAT	O35276	rattus norv
45	37.6	931	1	NRP2_HUMAN	O60462	homo sapien
46	37.6	931	1	NRP2_MOUSE	O35375	mus musculu
47	37.6	1021	1	S123_HUMAN	P55017	homo sapien
48	37.6	1805	1	NEST_RAT	P21363	rattus norv
49	37.6	1882	1	POL2_TRSVR	P25247	tomato ring
50	37.6	6632	1	UN89_CABEL	O01761	caenorhabdi
51	36.7	140	1	RS16_SCHPO	O60144	schizosacch
52	36.7	227	1	YMQ4_CABEL	P34495	caenorhabdi
53	36.7	275	1	YCT7_YEAST	P25627	saccharomyc
54	36.7	412	1	EFLH_YEAST	P36008	saccharomyc
55	36.7	435	1	RPOD_SORBI	Q01923	sorghum bic
56	36.7	466	1	MM13_RAT	P23097	rattus norv
57	36.7	747	1	VIVC_BPT7	P03725	bacterioph
58	36.7	1355	1	VG46_HSV11	Q00104	ictaluriid h
59	36.7	1527	1	RPOD_MAIZE	P16025	zea mays (m
60	36.7	3432	1	POLG_JAEV1	P27395	genome po
61	36.7	3432	1	POLG_JAEVS	P19110	genome po
62	36.7	3432	1	POLG_JAEVJ	P32886	genome po
63	35.8	127	1	PT02_STYPL	P28194	styeia plic
64	35.8	147	1	AI01_RAT	Q00104	ictaluriid h
65	35.8	196	1	KCY_MOUSE	P55009	rattus norv
66	35.8	222	1	RL4_CHLMU	Q9dbp5	mus musculu
67	35.8	259	1	ATE_RHIME	Q9bj15	chlamydia m
68	35.8	268	1	TRMB_MOUSE	Q92k18	rhizobium m
69	35.8	277	1	SPAS_DUNEI	Q92120	mus musculu
70	35.8	278	1	OSTP_SHEEP	P54214	dunaliella
71	35.8	310	1	DHLA_YANAU	Q9xxy9	ovis aries
72	35.8	353	1	ABR2_MOUSE	P22643	xanthobacte
73	35.8	401	1	PGK_HALVA	Q946j4	mus musculu
74	35.8	419	1	ENO_PYRAE	P50315	halocaula
75	35.8	482	1	MURE_THETN	Q8zye7	pyrobaculum
76	35.8	522	1	RF3_LISIN	Q8z9g2	thermoanaer
77	35.8	522	1	RF3_LISMO	Q92d33	listeria in
78	35.8	524	1	SBP_SOYBN	Q8v8c0	listeria mo
79	35.8	561	1	SR72_SCHPO	Q04672	glycine max
80	35.8	633	1	TDR4_HUMAN	Q59787	schizosacch
81	35.8	674	1	AMO_PEA	Q9nuy9	homo sapien
82	35.8	728	1	TRDN_HUMAN	Q13061	homo sapien
83	35.8	824	1	DAC1_XENLA	Q43077	pisum sativ
84	35.8	889	1	CHSC_ASPTU	Q8GG92	xenopus lae
85	35.8	1003	1	PHSL_VICFA	Q92197	aspergillus
86	35.8	1013	1	PTPX_MACNE	P53536	vicia faba
87	35.8	1051	1	PTK7_CHICK	O02695	macaca neme
88	35.8	1242	1	TBP7_CABEL	Q91048	gallus gall
89	35.3	373	1	DHVS_DIACA	P54816	caenorhabdi
90	35.3	417	1	GLYA_BUCAT	Q9axq8	dianthus ca
91	35.3	733	1	YC40_HUMAN	P57376	buchnera ap
92	35.3	1935	1	MYH7_PIG	Q9uli0	homo sapien
93	35.3	3924	1	ANK2_HUMAN	P79293	sus scrofa
94	35.3	78	1	YNP4_CABEL	Q01484	homo sapien
95	34.9	189	1	YPT2_CABEL	P34557	caenorhabdi
96	34.9	229	1	CSE4_YEAST	P41880	caenorhabdi
97	34.9	280	1	DHSX_MOUSE	P36012	saccharomyc
98	34.9	294	1	OSTP_MOUSE	Q8vbd0	mus musculu
99	34.9	303	1	OSTP_PIG	P10923	mus musculu
100	34.9	316	1	RLA0_PLAF8	P14287	sus scrofa
					Q94660	plasmodium

ALIGNMENTS

RESULT 1

POLG HPAV2 STANDARD; PRT; 2226 AA.

AC P26580;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]

DE P3D (EC 2.7.7.48)

OS Hepatitis A virus (strain 24a)

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12094;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=91162758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,

RA "Antigenic and genetic variation in cytopathic hepatitis A virus

RT variants arising during persistent infection: evidence for genetic

RT recombination."

RL J. Virol. 65:2056-2065 (1991).

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC [RNA] (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M59810; AAA45468.1; -.

DR MEROPS; C03.005; -.

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000605; RNA_helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00680; RNA_dep_RNA_pol_1.

DR Pfam; PF00910; RNA_helicase_1.

DR PRINTS; PR00918; CALICIVIRUSNS.

DR Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23

FT CHAIN 24 245

FT CHAIN 246 491

FT CHAIN 492 794

FT CHAIN 795 900

FT CHAIN 901 1087

FT CHAIN 1088 1422

FT CHAIN 1423 1495

FT CHAIN 1496 1518

FT CHAIN 1519 1737

FT CHAIN 1738 2226

SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 109; DB 1; Length 2226;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHTIECK 20

Db 779 VDDPRSEEDKRFESHTIECK 798

RESULT 2

POLG HPAV4 STANDARD; PRT; 2226 AA.

AC P26581;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]

DE P3D (EC 2.7.7.48)

OS Hepatitis A virus (strain 43c)

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12095;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=91162758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,

RA "Antigenic and genetic variation in cytopathic hepatitis A virus

RT variants arising during persistent infection: evidence for genetic

RT recombination."

RL J. Virol. 65:2056-2065 (1991).

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC [RNA] (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M59809; AAA45469.1; -.

DR MEROPS; C03.005; -.

DR InterPro; IPR004004; Calici_pol_hel.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000605; RNA_helicase.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR InterPro; IPR001205; RNA_pol_P3D.

DR InterPro; IPR007094; RNA_pol_PSVir.

DR InterPro; IPR008975; Viral_cap_coat.

DR Pfam; PF00680; RNA_dep_RNA_pol_1.

DR Pfam; PF00910; RNA_helicase_1.

DR PRINTS; PR00918; CALICIVIRUSNS.

DR Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.

FT CHAIN 1 23

FT CHAIN 24 245

FT CHAIN 246 491

FT CHAIN 492 794

FT CHAIN 795 900

FT CHAIN 901 1087

FT CHAIN 1088 1422

FT CHAIN 1423 1495

FT CHAIN 1496 1518

FT CHAIN 1519 1737

FT CHAIN 1738 2226

SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA90B9BF75 CRC64;

Query Match 100.0%; Score 109; DB 1; Length 2226;

Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHTIECK 20

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 3

POLG_HPAV8 STANDARD; PRT; 2226 AA.
AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065 (1991).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC [RNA] (N).
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59808; AAA45467.1; --
DR PDB; 1QA7; 15-MAY-00.
DR MEROPS; C03.005; --
DR InterPro; IPR004004; Calici pol hel.
DR InterPro; IPR009003; Cys Ser triypsin.
DR InterPro; IPR008605; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol P3D.
DR InterPro; IPR007094; RNA pol psvir.
DR InterPro; IPR008975; Viral cap coat.
DR Pfam; PF00680; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
FT CHAIN 2226 AA; 251292 MW; 24966463396C8D6B CRC64;
SQ SEQUENCE

Query Match 100.0%; Score 109; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20

Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4

POLG_HPAVH STANDARD; PRT; 2227 AA.
ID POLG_HPAVH
AC P08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501 (1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147 (1985).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M14114; AAA45475.1; --
CC EMBL; M14707; AAA45465.1; --
CC EMBL; M14707; AAA45466.1; ALT_INIT.
CC EMBL; M16632; AAA45471.1; --
CC PIR; A03905; A03905.
CC PIR; A25981; GNNYHW.

DR	PIR: A94149; GNNYMK.
DR	PDB; 1HAV; 23-DEC-96.
DR	MEROPOS; C03.C05; -.
DR	InterPro; IPR004004; Calici_pol_hel.
DR	InterPro; IPR005003; Cys Ser_trypsin.
DR	InterPro; IPR000603; RNA_helicase.
DR	InterPro; IPR007095; RNA_pol_DS_PS.
DR	InterPro; IPR001205; RNA_pol_P3D.
DR	InterPro; IPR007094; RNA_pol_PSwir.
DR	InterPro; IPR008975; Viral_cap_coat.
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR	Pfam; PF00910; RNA_helicase; 1.
DR	PRINTS; PR00918; CALICIVIRUSNS.
KW	Polypeptide; Coat protein; Core protein; Core protein; Transferase;
KW	RNA-directed RNA polymerase; Hydroxylase; Thiol processase; 3D-structure.
FT	CHAIN 1 23
FT	CHAIN 24 245
FT	CHAIN 246 491
FT	CHAIN 492 836
FT	CHAIN 837 980
FT	CHAIN 981 1087
FT	CHAIN 1088 1422
FT	CHAIN 1423 1496
FT	CHAIN 1497 1519
FT	CHAIN 1520 1738
FT	CHAIN 1739 2227
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT	CHAIN 77 77
FT	CHAIN 764 764
FT	CHAIN 821 821
FT	CHAIN 1052 1052
FT	CHAIN 1062 1062
FT	CHAIN 1118 1118
FT	CHAIN 1151 1151
FT	CHAIN 1163 1163
FT	CHAIN 1277 1277
FT	CHAIN 1500 1500
FT	CHAIN 1805 1805
FT	CHAIN 1930 1930
FT	CHAIN 2227 AA; 2251506 MW; 01E22957AE8740A6 CRC64;
FT</	

```

Query Match      100.0%; Score 109; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKFFESHIECRK 20
      |||||
Db 779 VDDPRSEEDKFFESHIECRK 798
      |||||

```

```

RESULT 5
POLG_HPAVM          STANDARD;          PRT;  2227 AA.
ID   POLG_HPAVM          STANDARD;          PRT;  2227 AA.
AC   P13901; Q81083; Q81084; Q81086; Q81087; Q81088; Q81089;
AC   Q81090; Q81091; Q81092; Q81093;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE   P2a TO P2c; Probable proteins P3a TO P3c; RNA-directed RNA polymerase
DE   P3d (EC 2.7.7.48)].
OS   Hepatitis A virus (strain MBB).
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Hepatocovirus.
NCBI_TaxID=12100;
[1]_RN
SEQUENCE FROM N.A.
RX   MEDLINE=88045071; PubMed=2823500;
RA   Paul A. V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
RA   Deinhardt F.;
RT   "The entire nucleotide sequence of the genome of human hepatitis A
RT   virus (isolate MBB).";
RL   Virus Res. 8:153-171(1987).
CC   -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

```


FT	CHAIN		1077	1422	CORE PROTEIN P2C.
FT	CHAIN		1423	1484	PROBABLE PROTEIN P3A.
FT	CHAIN		1485	1507	PROBABLE PROTEIN P3B.
FT	CHAIN		1508	1678	PROBABLE PROTEIN P3C.
FT	CHAIN		1679	2227	RNA-DIRECTED POLYMERASE P3D.
SQ	SEQUENCE	2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;			
 Query Match 97.2%; Score 106; DB 1; Length 2227; Best Local Similarity 95.0%; Pred. No. 4.9e-08; Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	VDDPRSEEDKRFESHIECK 20 :			
Ddb	779	VDDPRSEEDKRFESHIECK 798 :			
 RESULT 9 POLG HPVAVG ID POLG HPVAVG STANDARD; PRT; 808 AA. AC Q02381; AC 01-JUL-1993 (Rel. 26, Created) DT 01-JUL-1993 (Rel. 26, Last sequence update) DT 16-OCT-2001 (Rel. 40, Last annotation update) DT Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein DE P2A] (Fragment). DE Hepatitis A virus (strain GA76). OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; OC Hepatovirus. CC NCBI_TaxID=31706; CX [1] RN SEQUENCE FROM N.A. RP MEDLINE=92260183; PubMed=1316423; RA Khanna B., Speilbring J.E., Innis B.L., Robertson B.H.; RT "Characterization of a Genetic variant of human hepatitis A virus."; J. Med. Virol. 36:118-124(1992). CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4. CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins. ----- CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). ----- DR EMBL: M66695; AAA45477.1; -- DR InterPro: IPR008975; Viral_cap_coat. KW Polyprotein; Coat protein; Core protein. FT NON TER 1 1 FT CHAIN <1 2 COAT PROTEIN VP4 (PIA). FT CHAIN 3 223 COAT PROTEIN VP2 (PIB). FT CHAIN 224 470 COAT PROTEIN VP3 (PIC). FT CHAIN 471 770 COAT PROTEIN VP1 (PID). FT CHAIN 771 >808 CORE PROTEIN P2A. FT NON TER 808 808 SQ SEQUENCE 808 AA; 90632 MW; D80CE7E57AA479C12 CRC64;					
 Query Match 80.7%; Score 88; DB 1; Length 808; Best Local Similarity 80.0%; Pred. No. 9.4e-06; Matches 16; Conservative 3; Mismatches 1; Indels 0; Gaps 0;					
Qy	1	VDDPRSEEDKRFESHIECK 20 :			
Ddb	758	VDDPRTEEDKRFESHIECK 777 :			
 RESULT 10 POLG HPVAVS ID POLG HPVAVS STANDARD; PRT; 2230 AA. AC P14553;					

01-JAN-1990 (Rel. 13, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2G; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].

Simian hepatitis A virus (strain AGM-27).

Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.

NCBI_TaxID=12102;

[1]

SEQUENCE FROM N.A.

MEDLINE=913111420; PubMed=1649901;

Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R., Tsarev R.H.;

"Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";

J. Gen. Virol. 72:1677-1683(1991).

[2]

SEQUENCE OF 1750-2164 FROM N.A.

MEDLINE=89232168; PubMed=2541023;

Balayan M.S., Kusov Y.Y., Andapadize A.G., Tsarev S.A., Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;

"variations in genome fragments coding for RNA polymerase in human and simian hepatitis A viruses.";

FEBS Lett. 247:425-428(1989).

RNA} N}

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.

-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00924; BAA00766.1; -

EMBL; X15461; CAA33490.1; -

PIR; A30470; GNNYSA.

MEROPS; C03.005; -

InterPro; IPR004004; Calici_pol_hel.

InterPro; IPR009003; Cys_Ser_trypsin.

InterPro; IPR000605; RNA_helicase.

InterPro; IPR007095; RNA_pol_DS_Ps.

InterPro; IPR001205; RNA_pol_P3D.

InterPro; IPR007094; RNA_pol_PsVir.

InterPro; IPR008975; Viral_cap_coat.

Pfam; PF00680; RNA_dep_RNA_pol_1.

Pfam; PF00910; RNA_helicase; 1.

PRINTS; PR00918; CALICIVIRUSN.

Polyprotein; Coat protein; Core protein; Core protease; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

CHAIN 1 27 COAT PROTEIN VP4 (P1A).

CHAIN 28 249 COAT PROTEIN VP2 (P1B).

CHAIN 250 495 COAT PROTEIN VP3 (P1C).

CHAIN 496 795 COAT PROTEIN VP1 (P1D).

CHAIN 796 984 CORE PROTEIN P2A.

CHAIN 985 1091 CORE PROTEIN P2B.

CHAIN 1092 1426 CORE PROTEIN P2C.

CHAIN 1427 1498 PROBABLE PROTEIN 3A.

CHAIN 1499 1521 PROBABLE PROTEIN 3B.

CHAIN 1522 1741 PROBABLE PROTEIN 3C.

CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.

SEQUENCE 2230 AA; 251296 MW; 87B3230B324E1F19 CRC64;

Query Match 77.1%; Score 84; DB 1; Length 2230;

Best Local Similarity 94.1%; Pred. No. 0.00012;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
 |||||:|||||
 DB 783 VDDPRSEEDKRFESHIE 799

RESULT 11

POLG_HPAVT
 ID POLG_HPAVT STANDARD; PRT; 839 AA.
 AC P31788;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 DE P2A)
 OS Simian hepatitis A virus (strain CV-145).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=31707;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=91311421; PubMed=1649902;
 RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
 RT "Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (Macaca fascicularis).";
 RL J. Gen. Virol. 72:1685-1689(1991).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M59286; AAA45473.1; -;
 DR PIR; JQ1180; GNNYS2.
 DR InterPro; IPR008975; Viral_cap_cat.
 KW Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 ? COAT PROTEIN VP1 (P1D).
 FT CHAIN ? >839 CORE PROTEIN P2A.
 FT NON_TER 839 839
 SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;

Query Match 74.3%; Score 81; DB 1; Length 839;
 Best Local Similarity 82.4%; Pred. No. 0.00012;
 Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
 |||||:|||||
 DB 779 VDDPRTEDEDRRFESHIE 795

RESULT 12

CO3_RAT
 ID CO3_RAT STANDARD; PRT; 1663 AA.
 AC P01026;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Complement C3 precursor [Contains: C3a anaphylatoxin].
 DE C3.
 GN Rattus norvegicus (Rat).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Liver;
 RX MEDLINE=90245672; PubMed=2336397;
 RA Misumi Y., Sohma M., Ikehara Y.;
 RT "Nucleotide and deduced amino acid sequence of rat complement C3.";
 RL Nucleic Acids Res. 18:2178-2178(1990).
 RN [2]
 RP SEQUENCE OF 671-748.
 RX MEDLINE=79062262; PubMed=309768;
 RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
 RA Daniels J.S., Daughaday W.H., Bradshaw R.A.;
 RT "Purification, characterization, and amino acid sequence of rat anaphylatoxin (C3a).";
 RL Biochemistry 17:5031-5038(1978).
 RN [3]
 RP SEQUENCE OF 1316-1595 FROM N.A.
 RX MEDLINE=89380332; PubMed=2674144;
 RA Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
 RA Lyttle C.R.;
 RT "Estrogen regulation of tissue-specific expression of complement C3";
 RL J. Biol. Chem. 264:16941-16947(1989).
 CC -!- FUNCTION: C3 plays a central role in the activation of the complement system. Its processing by C3 convertase is the central reaction in both classical and alternative complement pathways. After activation C3b can bind covalently, via its reactive thioester, to cell surface carbohydrates or immune aggregates.
 CC -!- FUNCTION: Derived from proteolytic degradation of complement C3, C3a anaphylatoxin is a mediator of local inflammatory process. It induces the contraction of smooth muscle, increases vascular permeability and causes histamine release from mast cells and basophilic leukocytes.
 CC -!- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN, RELEASING C3a ANAPHYLATOXIN & GENERATING C3b (BETA CHAIN + ALPHA CHAIN).
 CC -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
 CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
 CC -!- SIMILARITY: Contains 1 NTR domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52477; CAA36716.1; -;
 DR EMBL; M29866; AAA40837.1; ALT_SEQ.
 DR PIR; S15764; C3RT.
 DR PDB; 1QQF; 31-JUL-00.
 DR PDB; 1QSU; 31-JUL-00.
 DR InterPro; IPR002890; A2M_N.
 DR InterPro; IPR009048; AM_receptor_bind.
 DR InterPro; IPR000020; Anaphylatoxin.
 DR InterPro; IPR001840; Anaphylatoxin.
 DR InterPro; IPR001599; MacrogloblnA2.
 DR InterPro; IPR001134; Netrin_C.
 DR InterPro; IPR008930; Terp_cyc_toroid.
 DR Pfam; PF00207; A2M; 1.
 DR Pfam; PF01835; A2M_N; 1.
 DR Pfam; PF01821; ANATO; 1.
 DR Pfam; PF01759; NTR; 1.
 DR PRINTS; PR000004; ANAPHYLATOXN.
 DR PRODOM; PD003264; Anaphylatoxin; 1.
 DR SMART; SM00104; ANATO; 1.
 DR SMART; SM00643; C345C; 1.
 DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
 DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
 DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.

complement system. Its processing by C3 convertase is the central reaction in both classical and alternative complement pathways. After activation C3b can bind covalently, via its reactive thioester, to cell surface carbohydrates or immune aggregates. C3a anaphylatoxin is a mediator of local inflammatory process. It induces the contraction of smooth muscle, increases vascular permeability and causes histamine release from mast cells and basophilic leukocytes.

- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg residues, forming two chains, beta and alpha, linked by a disulfide bond. C3 convertase activates C3 by cleaving the alpha chain, releasing C3a anaphylatoxin and generating C3b (beta chain + alpha' chain).

- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.

- SIMILARITY: Contains 1 anaphylatoxin-like domain.

- SIMILARITY: Contains 1 NTR domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; M34054; AAA37038.1; --
PIR; A37156; C3GP.
HSSP; P01024; 1C3D.
InterPro; IPR002890; A2M_N.
InterPro; IPR009048; AM_Receptor_bind.
InterPro; IPR000020; Anaphylatoxin.
InterPro; IPR001840; Anaphylatoxin.
InterPro; IPR001599; MacrogloblnA2.
InterPro; IPR001134; Netrin-C.
InterPro; IPR008930; Terp_cyc_toroid.
Pfam; PF00207; A2M; 1.
Pfam; PF01835; A2M_N; 1.
Pfam; PF01821; ANATO; 1.
Pfam; PF01759; NTR; 1.
PRINTS; PR00004; ANAPHYLATOXN.
ProDom; PD003264; Anaphylatoxin; 1.
SMART; SM00104; ANATO; 1.
SMART; SM00643; C345C; 1.
PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
PROSITE; PS01177; ANAPHYLATOXIN_1; 1.
PROSITE; PS01178; ANAPHYLATOXIN_2; 1.
PROSITE; PS01189; NTR; 1.
Complement pathway; Complement alternate pathway; Plasma; Inflammatory response; Glycoprotein; Signal; Thioester bond.

SIGNAL 1 29
CHAIN 30 1666 COMPLEMENT C3.
CHAIN 30 671 BETA CHAIN.
CHAIN 676 1666 ALPHA CHAIN.
PEPTIDE 676 753 C3A ANAPHYLATOXIN.
CHAIN 754 1666 C3B (ALPHA' CHAIN).
DOMAIN 698 733 ANAPHYLATOXIN-LIKE.
DOMAIN 1522 1664 NTR.
SITE 753 754 CLEAVAGE (BY C3 CONVERTASE).
DISULFID 557 821 INTERCHAIN (BY SIMILARITY).
DISULFID 630 666 BY SIMILARITY.
DISULFID 698 725 BY SIMILARITY.
DISULFID 699 732 BY SIMILARITY.
DISULFID 712 733 BY SIMILARITY.
DISULFID 878 1517 BY SIMILARITY.
DISULFID 1106 1163 BY SIMILARITY.
DISULFID 1363 1493 BY SIMILARITY.
DISULFID 1394 1462 BY SIMILARITY.
DISULFID 1510 1515 BY SIMILARITY.
DISULFID 1522 1593 BY SIMILARITY.
DISULFID 1540 1664 BY SIMILARITY.
DISULFID 1640 1649 BY SIMILARITY.
CROSSLINK 1015 1018 Iso-glutamyl cysteine thioester (Cys-Gln).

PROSITE; PS0189; NTR; 1.
Complement pathway; Complement alternate pathway; Plasma; Inflammatory response; Glycoprotein; Signal; 3D-structure; Thioester bond.

SIGNAL 1 24
CHAIN 25 1663 COMPLEMENT C3.
CHAIN 25 666 BETA CHAIN.
CHAIN 671 1663 ALPHA CHAIN.
PEPTIDE 671 748 C3A ANAPHYLATOXIN.
CHAIN 749 1663 C3B (ALPHA' CHAIN).
DOMAIN 693 728 ANAPHYLATOXIN-LIKE.
DOMAIN 1518 1661 NTR.
SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).
DISULFID 558 816 INTERCHAIN (BY SIMILARITY).
DISULFID 626 661 BY SIMILARITY.
DISULFID 693 720 BY SIMILARITY.
DISULFID 694 727 BY SIMILARITY.
DISULFID 707 728 BY SIMILARITY.
DISULFID 873 1513 BY SIMILARITY.
DISULFID 1101 1158 BY SIMILARITY.
DISULFID 1358 1489 BY SIMILARITY.
DISULFID 1389 1458 BY SIMILARITY.
DISULFID 1506 1511 BY SIMILARITY.
DISULFID 1518 1590 BY SIMILARITY.
DISULFID 1537 1661 BY SIMILARITY.
CROSSLINK 1010 1013 Iso-glutamyl cysteine thioester (Cys-Gln).
CARBOHYD 939 939 N-LINKED (GLCNAC. .) (PROBABLE).
CARBOHYD 1617 1617 N-LINKED (GLCNAC. .) (PROBABLE).
CONFLICT 721 722 LK -> XL (IN REF. 2).
SEQUENCE 1663 AA; 18666 MW; 2F87CCB143CDD4BC CRC64;

Query Match 45.9%; Score 50; DB 1; Length 1663;
Best Local Similarity 38.9%; Pred. No. 14;
Matches 7; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 2 DDPREEDKRFESHICR 19
1574 DEVQAGERRFISHVKCR 1591

RESULT 13
CO3_CAVPO STANDARD; PRT; 1666 AA.
ID_CO3_CAVPO
AC P12387; 1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 43, Last annotation update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Complement C3 precursor [Contains: C3a anaphylatoxin].
GN C3.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90307998; PubMed=19731176;
RA Auerbach H.S., Burger R., Dodds A., Colten H.R.;
RT "Molecular basis of complement C3 deficiency in guinea pigs.";
RL J. Clin. Invest. 86:96-106(1990).
RN [2]
RP SEQUENCE OF 676-753.
RX MEDLINE=89113342; PubMed=3064079;
RA Gerard N.P., Lively M.O., Gerard C.;
RT "Amino acid sequence of guinea pig C3a anaphylatoxin.";
RL protein Seq. Data Anal. 1:473-478(1988).
RN [3]
RP SEQUENCE OF 993-1032.
RX MEDLINE=89178889; PubMed=6838833;
RA Thomas M.L., Tack B.F.;
RT "Identification and alignment of a thiol ester site in the third component of guinea pig complement.";
RL Biochemistry 22:942-947(1983).
RN [4]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [5]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [6]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [7]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [8]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [9]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [10]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [11]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [12]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [13]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [14]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [15]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [16]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [17]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [18]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [19]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [20]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [21]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [22]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [23]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [24]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [25]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [26]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [27]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [28]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [29]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [30]
RP Biochemistry 22:942-947(1983).
RL Biochemistry 22:942-947(1983).
RN [

FT NON TER 1
DOMAIN 177 321 NTR.
SEQUENCE 323 AA; 37070 MW; 3C8011BD9A1F15AD CRC64;

Query Match 44.0%; Score 48; DB 1; Length 323;
Best Local Similarity 42.1%; Pred.No. 4.8;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECRK 20
| : : : ||||
DB 234 DEPDCKTRNFISHIKRK 252

RESULT 15
Y9E3 PSEA
ID Y9E3 PSEA STANDARD; ERT; 242 AA.
AC P33663;
DT DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical UPF0124 protein PA4543.
GN PA4543.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_taxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=2043737; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P., Hickey M.J., Brinkman F.S.L., Huffnagle W.O., Kowalik D.J., Lagrou M., Garber R.B., Coltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen.";
RT Nature 406:959-964(2000).
RL [2]

RN RL
RP SEQUENCE OF 198-242 FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=93225810; PubMed=8097014;
RA Hobbs M., Collie E.S., Free P.D., Livingston S.P., Mattick J.S.;
RA "Pils and PilR, a two-component transcriptional regulatory system controlling expression of type 4 fimbriae in Pseudomonas aeruginosa";
RT Mol. Microbiol. 7:669-682(1993).
RL CC
CC -!- SIMILARITY: Belongs to the UPF0124 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).

EMBL; AE0004868; ARG07931.1; -.
DR EMBL; LO6013; AAA87639.1; -.
DR PIR; B83077; E83077.
DR InterPro; IPR003730; DUF152.
DR Pfam; PF02578; DUF152; 1.
DR TIGRFAMs; TIGR00726; TIGR00726; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 242 AA; 26047 MW; AD4470FF676FF51D CRC64;

Query Match 42.2%; Score 46; DB 1; Length 242;
Best Local Similarity 45.0%; Pred.No. 7.1;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

QY 2 DDPRSEED--KRPFESHTICR 19
||||| : : |:::

Db	42	DDPRAVEENRRRLTERLECR	61	
RESULT 16				
CO3 MOUSE		STANDARD;	PRT;	1663 AA.
AD	CO3 MOUSE			P01027;
DT	21-JUL-1986	(Rel. 01. Created)		
DT	01-FEB-1986	(Rel. 33, Last sequence update)		
DT	15-MAR-2004	(Rel. 43, Last annotation update)		
DE	Complement C3 precursor (HSE-MSF) [Contains: C3a anaphylatoxin].			
GN	C3.			
OS	Mus musculus (Mouse).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM LONG).			
RX	MEDLINE=8503854; PubMed=6208565;			
RA	Fey G.H., Lundwall A., Wetsel R.A., Tack B.F., de Bruijn M.H.L.,			
RA	Domdey H.;			
RT	"Nucleotide sequence of complementary DNA and derived amino acid			
RT	sequence of murine complement protein C3."			
RL	Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:333-344(1984).			
RN	[2]			
RP	SEQUENCE OF 671-1663 FROM N.A. (ISOFORM LONG).			
RX	MEDLINE=85054819; PubMed=6094532;			
RA	Wetsel R.A., Lundwall A., Davidson F., Gibson T., Tack B.F., Fey G.H.;			
RT	"Structure of murine complement component C3. II. Nucleotide sequence			
RT	of cloned complementary DNA coding for the alpha chain."			
RL	J. Biol. Chem. 259:13857-13862(1984).			
RN	[3]			
RP	SEQUENCE OF 671-748 FROM N.A.			
RX	MEDLINE=83117730; PubMed=6961437;			
RA	Domdey H., Wiebauer K., Kazmaier M., Mueller V., Odink K., Fey G.H.;			
RT	"Characterization of the mRNA and cloned cDNA specifying the third			
RT	component of mouse complement."			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:7619-7623(1982).			
RN	[4]			
RP	SEQUENCE OF 658-761 FROM N.A.			
RX	MEDLINE=84201365; PubMed=6609661;			
RA	Fey G.H., Wiebauer K., Domdey H.;			
RT	"Amino acid sequences of mouse complement C3 derived from nucleotide			
RT	sequences of cloned cDNA."			
RL	Ann. N.Y. Acad. Sci. 421:307-312(1983).			
RN	[5]			
RP	SEQUENCE OF 1-34 FROM N.A.			
RX	MEDLINE=83117622; PubMed=6985486;			
RA	Wiebauer K., Domdey H., Digelmann H., Fey G.;			
RT	"Isolation and analysis of genomic DNA clones encoding the third			
RT	component of mouse complement."			
RL	Proc. Natl. Acad. Sci. U.S.A. 79:7077-7081(1982).			
RN	[6]			
RP	SEQUENCE OF 25-41 AND 749-760.			
RX	MEDLINE=93373334; PubMed=8364938;			
RA	Hamada J.-I., Cavanaugh P.G., Miki K., Nicolson G.L.;			
RT	"A paracrine migration-stimulating factor for metastatic tumor cells			
RT	secreted by mouse hepatic sinusoidal endothelial cells:			
RT	identification as complement component C3b."			
RL	Cancer Res. 53:4418-4423(1993).			
RN	[7]			
RP	ALTERNATIVE INITIATION.			
RX	MEDLINE=95053742; PubMed=7964485;			
RA	Cahen-Kramer Y., Martenson I.L., Melchers F.;			
RT	"The structure of an alternate form of complement C3 that displays			
RT	costimulatory growth factor activity for B lymphocytes."			
RL	J. Exp. Med. 180:2079-2088(1994).			
CC	-1- FUNCTION: C3 plays a central role in the activation of the			
CC	complement system. Its processing by C3 convertase is the central			
CC	reaction in both classical and alternative complement pathways.			
CC	After activation C3b can bind covalently, via its reactive			
CC	thioester, to cell surface carbohydrates or immune aggregates.			
CC	-1- FUNCTION: Derived from proteolytic degradation of complement C3,			
CC				
CC	C3a anaphylatoxin is a mediator of local inflammatory process. It			
CC	induces the contraction of smooth muscle, increases vascular			
CC	permeability and causes histamine release from mast cells and			
CC	basophilic leukocytes. The short isoform has B-cell stimulatory			
CC	activity.			
CC	-1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg			
CC	residues, forming two chains, beta and alpha, linked by a			
CC	disulfide bond. C3 convertase activates C3 by cleaving the alpha			
CC	chain, releasing C3a anaphylatoxin and generating C3b (beta chain			
CC	+ alpha' chain)			
CC	-1- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative initiation;			
CC	Comment=2 isoforms, Long (shown here) and Short, are produced by			
CC	alternative initiation;			
CC	-1- PTM: C3b is rapidly split in two positions by factor I and a			
CC	cofactor to form iC3b (inactivated C3b) and C3f which is released.			
CC	Then iC3b is slowly cleaved (possibly by factor I) to form C3c			
CC	and C3dg. Other proteases produce other fragments such as C3d or			
CC	C3g.			
CC	-1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.			
CC	-1- SIMILARITY: Contains 1 anaphylatoxin-like domain.			
CC	-1- SIMILARITY: Contains 1 NTR domain.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; K02782; AAC42013.1; -			
DR	EMBL; J00369; AAA37336.1; -			
DR	EMBL; J00367; AAA37336.1; JOINED.			
DR	EMBL; M33032; AAA37378.1; -			
DR	EMBL; Z37998; CAA86099.2; -			
DR	PIR; A92459; C3MS.			
DR	HSSP; P01024; 1C3D.			
DR	MGD; MGI:88227; C3.			
DR	InterPro; IPR002890; A2M_N.			
DR	InterPro; IPR009048; AM_receptor_bind.			
DR	InterPro; IPR000020; Anaphylatoxin.			
DR	InterPro; IPR001840; Anaphylatoxn.			
DR	InterPro; IPR001599; MacroglobinA2.			
DR	InterPro; IPR001134; Netrin_C.			
DR	InterPro; IPR008930; Terp_cyc_toroid.			
DR	Pfam; PF00207; A2M; 1.			
DR	Pfam; PF01835; A2M_N; 1.			
DR	Pfam; PF01821; ANATO; 1.			
DR	Pfam; PF01759; NTR; 1.			
DR	PRINTS; PR00004; ANAPHYLATOXN.			
DR	ProDom; PD003264; Anaphylatoxin; 1.			
DR	SMART; SM00104; ANATO; 1.			
DR	SMART; SM00643; C345C; 1.			
DR	PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.			
DR	PROSITE; PS01177; ANAPHYLATOXIN_1; 1.			
DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 1.			
DR	PROSITE; PS0189; NTR; 1.			
KW	Complement pathway; Complement alternate pathway; Plasma;			
KW	Inflammatory response; Glycoprotein; Signal; Alternative initiation;			
KW	Thioester bond.			
FT	SIGNAL	1	24	
FT	CHAIN	25	1663	COMPLEMENT C3, ISOFORM LONG.
FT	CHAIN	25	666	COMPLEMENT C3, BETA CHAIN.
FT	CHAIN	671	1663	COMPLEMENT C3, ALPHA CHAIN.
FT	CHAIN	1129	1663	COMPLEMENT C3, ISOFORM SHORT.
FT	INIT_MET	1129	1129	FOR ISOFORM SHORT.
FT	PEPTIDE	671	748	C3A ANAPHYLATOXIN.
FT	CHAIN	749	1663	C3B (ALPHA' CHAIN).
FT	PEPTIDE	749	954	C3C FRAGMENT.
FT	PEPTIDE	955	1303	C3DG FRAGMENT.
FT	PEPTIDE	955	1001	C3G FRAGMENT.
FT	PEPTIDE	1002	1303	C3D FRAGMENT.

PEPTIDE 1304 1320 C3F FRAGMENT.
FT DOMAIN 693 728 ANAPHYLATOXIN-LIKE.
FT DOMAIN 1518 1661 NTR.
FT SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).
FT SITE 1303 1304 CLEAVAGE (BY FACTOR I).
FT SITE 1320 1321 CLEAVAGE (BY FACTOR I).
FT SITE 559 616 INTERCHAIN (BY SIMILARITY).
FT DISULFID 626 661 BY SIMILARITY.
FT DISULFID 693 720 BY SIMILARITY.
FT DISULFID 694 727 BY SIMILARITY.
FT DISULFID 707 728 BY SIMILARITY.
FT DISULFID 873 1513 BY SIMILARITY.
FT DISULFID 1101 1158 BY SIMILARITY.
FT DISULFID 1358 1489 BY SIMILARITY.
FT DISULFID 1389 1458 BY SIMILARITY.
FT DISULFID 1506 1511 BY SIMILARITY.
FT DISULFID 1518 1590 BY SIMILARITY.
FT DISULFID 1537 1661 BY SIMILARITY.
FT DISULFID 1637 1646 BY SIMILARITY.
FT CARBOHYD 939 939 N-LINKED (GLCNAC. . .).
FT CARBOHYD 1617 1617 N-LINKED (GLCNAC. . .).
FT CROSSLINK 1010 1013 Isoglutamyl cysteine thioester (Cys-Gln) (By similarity).
SQ SEQUENCE 1663 AA; 186482 MW; DE5546CC769BEA19 CRC64;
Query Match 42.2%; Score 46; DB 1; Length 1663;
Best Local Similarity 38.9%; Pred. No. 58;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;
QY 2 DDPRSEEDKRFESHIECR 19
DB 1574 DEVQAGQQRKFISHIKR 1591
RESULT 17
YD35 MYCPN
ID YD35 MYCPN STANDARD; PRT; 741 AA.
AC P75443;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein MPN335 (F10_orf741).
GN MPN335 OR MP501.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae";
RL Nucleic Acids Res. 24:4420-4449(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO M.PNEUMONIAE MPN333.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; AE000049; AAB96149.1; -.
DR PIR; S73827; S73827.
DR KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 120 140 POTENTIAL.
FT TRANSMEM 156 176 POTENTIAL.

FT TRANSMEM 187 207 POTENTIAL.
FT TRANSMEM 655 675 POTENTIAL.
FT TRANSMEM 715 735 POTENTIAL.
SQ SEQUENCE 741 AA; 85025 MW; 1776A96BCF83567F CRC64;
Query Match 41.3%; Score 45; DB 1; Length 741;
Best Local Similarity 42.1%; Pred. No. 34;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 2 DDPRSEEDKRFESHIECR 20
DB 629 EDKQSEEEKFDEIENAK 647
RESULT 18
RL4 CHLTR
ID RL4 CHLTR STANDARD; PRT; 222 AA.
AC O84532;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L4.
GN RPLD OR RL4 OR CT527.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=9900809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V., Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis";
RL Science 282:754-759(1998).
CC -1- FUNCTION: This protein binds directly and specifically to 23S rRNA (By similarity).
CC -1- SIMILARITY: Belongs to the L4P family of ribosomal proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; AE001323; AAC68128.1; -.
DR PIR; E71507; E71507.
DR InterPro; IPR002136; Ribosomal_L4/L1E.
DR Pfam; PF00573; Ribosomal_L4; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 222 AA; 24506 MW; 44641751527630D1 CRC64;
Query Match 40.4%; Score 44; DB 1; Length 222;
Best Local Similarity 42.9%; Pred. No. 13;
Matches 9; Conservative 6; Mismatches 4; Indels 2; Gaps 1;
QY 1 VDDPRSEEDKRF--ESHIECR 19
DB 135 LDAPKTKEARFLKCNVECR 155
RESULT 19
TRMB_HUMAN
ID TRMB_HUMAN STANDARD; PRT; 276 AA.
AC Q9UBP6; Q14105;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE tRNA (guanine-N(7)-)-methyltransferase (EC 2.1.1.33) (tRNA(m7G46)-methyltransferase) (Methyltransferase-like protein 1).

GN METTL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC
RX MEDLINE=99263505; PubMed=10329009;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C., Rodriguez A.C., Grinnow J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP FUNCTION, AND INTERACTION WITH WDR4.
RX MEDLINE=22290633; PubMed=12403464;
RA Alexandrov A., Martzen M.R., Phizicky E.M.;
RT "Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA.";
RL RNA 8:1253-1256(2002).
CC -!- FUNCTION: Catalyzes the formation of N(7)-methylguanine at position 46 (m7G46) in tRNA.
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-homocysteine + tRNA containing N(7)-methylguanine.
CC -!- SUBUNIT: Forms a complex with WDR4.
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. Trmb family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: Y18643; CAA77239.1; --
CC EMBL: Y18642; CAA77238.1; --
CC EMBL: X96698; CAA65470.1; ALT_INIT.
CC EMBL: BC000550; AAH00550.1; --
CC Genew: HGNC:7030; METTL1.
CC MIM: 604465; --
CC GO: GO:0008169; F:methyltransferase activity; TAS.
CC GO: GO:0008176; F:tRNA (guanine-N7-)-methyltransferase activity; IDA.
CC GO: GO:0006400; P:tRNA modification; IDA.
CC InterPro: IPR004395; Cons hypoth91.
CC InterPro: IPR003358; Methyltransf_4.
CC InterPro: IPR000051; SAM_bind.

DR Pfam: PF02390; Methyltransf 4; 1.
KW TIGRFAW; TIGR00091; TIGR00091; 1.
KW Transferase; Methyltransferase; LRNA processing.
SQ SEQUENCE 276 AA; 31471 MW; 715AE65A18632892 CRC64;
Query Match 40.4%; Score 44; DB 1; Length 276;
Best Local Similarity 37.5%; Pred. No. 17;
Matches 6; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
QY 2 DDPSEEDKRPESHIE 17
DB 64 DDPKDKKRAQAQVE 79
RESULT 20
FURC_DROME STANDARD; PRT; 1101 AA.
ID FURC_DROME STANDARD; PRT; 1101 AA.
AC P30430; Q27235; Q9VBR5;
DT 01-APR-1993 (Rel. 25, Created)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Furin-like protease 1, isoform 1-CRR precursor (EC 3.4.21.75) (Furin 1) (Kex2-like endoprotease 1) (dKLP-1).
GN FUR1 OR CGI0772.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Head;
RX MEDLINE=92185516; PubMed=1545235;
RA Hayflick J.S., Wolfgang W.J., Forte M.A., Thomas G.;
RT "A unique Kex2-like endoprotease from Drosophila melanogaster is expressed in the central nervous system during early embryogenesis.";
RL J. Neurosci. 12:705-717(1992).
RN [2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
RC STRAIN=Tuebingen, and Oregon-R; TISSUE=Embryo;
RX MEDLINE=93259127; PubMed=8491178;
RA Roebroek A.J.M., Cremers J.W.M., Pauli I.G.L., Bogaert T., Van de Ven W.J.M.;
RT "Generation of structural and functional diversity in furin-like proteins in Drosophila melanogaster by alternative splicing of the Dfurl gene.";
RL EMBL J. 12:1853-1870(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Ananides P.G., Scherer S.E., Li P.W., Hosking R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Bason K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadeu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Flesler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svarksas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Weissman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
RN [4]
RP REVISIONS, AND ALTERNATIVE SPLICING.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S., Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E., Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P., Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M., Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Furin is likely to represent the ubiquitous endoprotease activity within constitutive secretory pathways and capable of cleavage at the RX(K/R)R consensus motif (by similarity).
CC -!- CATALYTIC ACTIVITY: Release of mature proteins from their propeptides by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa can be any amino acid and Yaa is Arg or Lys. Releases albumin, complement component C3 and von Willebrand factor from their respective precursors.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Golgi.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Comment-Experimental confirmation may be lacking for some isoforms;
CC Name=1-CRR; Synonyms=A;
CC IsoId=P30430-1; Sequence=Displayed;
CC Name=1; Synonyms=F;
CC IsoId=P26016-2; Sequence=External;
CC Name=1-X; Synonyms=B;
CC IsoId=P26016-1; Sequence=External;
CC Name=2; Synonyms=C, D, E;
CC IsoId=P26016-3; Sequence=External;
CC -!- TISSUE SPECIFICITY: In adults, isoform 1-CRR is expressed in CNS, fat body, and female reproductive tissues, and in embryos, in anal pads, hindgut, developing antennomaxillary complex, oenocytes, clypeolabrum, pharynx, trachea, CNS and developing posterior spiracles.
CC -!- DEVELOPMENTAL STAGE: Isoform 1-CRR is expressed in embryos, larvae, pupae and adults.
CC -!- SIMILARITY: Belongs to peptidase family S8. Furin subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M81431; AAA28467.1; -
CC EMBL; L12370; AAA28545.1; -
CC EMBL; L12369; AAA28545.1; JOINED.
CC EMBL; L12372; AAA28548.1; -
CC EMBL; AB003751; AANI4051.1; -

DR MEROPS; S08.048; -
DR FlyBase; FBgn0004509; Furl.
DR GO; GO:0005886; C:plasma membrane; IDA.
DR GO; GO:0004276; F:furin activity; IDA.
DR InterPro; IPR006212; Furin_repeat.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR002884; Peptidase_S8B.
DR InterPro; IPR009020; Protease_inhib.
DR Pfam; PF01483; P_protease; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR PRODOM; PD000717; P_domain; 1.
DR SMART; SMC0261; FU; 2.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Signal; Transmembrane;
KW Multigene family; Zymogen; Alternative splicing; Golgi stack.
FT SIGNAL 1 ? POTENTIAL.
FT PROPEP ? 309 POTENTIAL.
FT CHAIN 310 1101 FURIN-LIKE PROTEASE 1, ISOFORM 1-CRR.
FT DOMAIN 310 655 CATALYTIC.
FT ACT_SITE 372 372 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 413 413 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 587 587 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT TRANSMEM 119 139 POTENTIAL.
FT DOMAIN 1014 1034 POTENTIAL.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 108 108 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 426 426 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 727 727 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 859 859 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 978 978 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 1014 1014 T -> I (IN REF. 1).
SQ SEQUENCE 1101 AA; 120993 MW; 90DC38E2CACB71A0 CRC64;
Query Match 40.4%; Score 44; DB 1; Length 1101;
Best Local Similarity 42.1%; Pred. No. 75;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 2 DDPRESSEDRFESHIECRK 20
Db 1057 DESDSEDELFTHFPARK 1075
RESULT 21
ID DHGA ACICA STANDARD; PRT; 801 AA.
AC P03465;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose dehydrogenase-A [pyrroloquinoline-quinone] precursor
DE (EC 1.1.99.17) (Quinoprotein glucose DH) (GDH-A).
GN GDHA.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Moraxellaceae; Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LMD 79.41;
RX MEDLINE=88289368; PubMed=3399393;
RA Clifton-Jensen A.-M., Goosen N., Odle G., van de Putte P.;
RT "Nucleotide sequence of the gene coding for quinoprotein glucose dehydrogenase from Acinetobacter calcoaceticus";
RL Nucleic Acids Res. 16:6228-6228(1988).
CC -!- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +

reduced acceptor.
 CC -1- COFACTOR: PQQ.
 CC -1- SUBUNIT: Monomer.
 CC -1- SUBCELLULAR LOCATION: Periplasmic.
 CC -1- MISCELLANEOUS: ACINETOBACTER CALCOACETICUS CONTAINS TWO DIFFERENT
 CC PQQ DEPENDENT ENZYMES WITH GDH ACTIVITY. GDH-A PREFERENCES
 CC 2-DEOXYGLUCOSE AS SUBSTRATE, THE SPECIFIC SUBSTRATES FOR GDH-B ARE
 CC DISACCHARIDES.
 CC -1- SIMILARITY: Belongs to the bacterial PQQ dehydrogenase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X07235; CAA30222.1; -.
 CC PIR; S00943; S00943.
 CC InterPro; IPR001479; Bac_PQQ.
 CC InterPro; IPR002372; Bac_PQQ_repeat.
 CC Pfam; PF01011; PQQ; 4.
 CC SMART; SM00564; PQQ; 5.
 CC PROSITE; PS00363; BACTERIAL_PQQ_1; 1.
 CC PROSITE; PS00364; BACTERIAL_PQQ_2; 1.
 CC Oxidoreductase; PQQ; Transmembrane; Periplasmic; signal.
 CC SIGNAL 1 33 POTENTIAL.
 CC CHAIN 34 801
 CC [PYRROLOQUINOLINE-QUINONE].
 CC TRANSMEM 39 55
 CC TRANSMEM 59 79 POTENTIAL.
 CC TRANSMEM 94 108 POTENTIAL.
 CC TRANSMEM 119 138 POTENTIAL.
 CC ACT_SITE 471 471 BASE (POTENTIAL).
 CC SEQUENCE 801 AA; 86956 MW; 2F4EA22EB5450D8 CRC64;
 SQ
 Query Match 39.9%; Score 43.5; DB 1; Length 801;
 Best Local Similarity 47.1%; Pred. No. 63;
 Matches 8; Conservative 4; Mismatches 4; Indels 1; Gaps 1;
 QY 3 DPASEEDKRFESHIECR 19
 Db 249 DPKLTKDSFQ-HLTCR 264
 RESULT 22
 GCTB_ACIFE STANDARD; PRT; 265 AA.
 AC Q59112;
 DT 15-JUL-1998 (rel. 36, Created)
 DT 30-MAY-2000 (rel. 39, Last sequence update)
 DT 28-FEB-2003 (rel. 41, Last annotation update)
 DE Glutacate CoA-transferase subunit B (EC 2.8.3.12) (GCT small
 DE subunit).
 GN GCTB.
 OS Acidaminococcus fermentans.
 OS Bacteria; Firmicutes; Clostridia; Clostridiales; Acidaminococcaceae;
 OC Acidaminococcus.
 OC NCBI_taxid=905;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18 AND 43-65.
 RC STRAIN=ATCC 25085;
 RX MEDLINE=95045599; PubMed=7957258;
 RA Mack M., Bendrat K., Zelder O., Eckel E., Linder D., Buckel W.;
 RT "Location of the two genes encoding glutacate coenzyme
 RT A-transferase at the beginning of the hydroxyglutarate operon in
 RT Acidaminococcus fermentans."
 RL Eur. J. Biochem. 226:41-51(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS).
 RC STRAIN=ATCC 25085;
 RX MEDLINE=97238937; PubMed=9083111;

Jacob U., Mack M., Clausen T., Huber R., Buckel W., Messerschmidt A.;
 "Glutacate CoA-transferase from Acidaminococcus fermentans: the
 crystal structure reveals homology with other CoA-transferases";
 Structure 5:415-426(1997).
 CC -1- FUNCTION: CATALYZES THE TRANSFER OF THE COA MOIETY FROM ACETYL COA
 CC TO (R)-2-HYDROXYGLUTARATE AND RELATED COMPOUNDS LIKE GLUTACONATE.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + (E)-glutacate = acetate +
 CC glutacetyl-1-CoA.
 CC -1- PATHWAY: Glutamate fermentation.
 CC -1- SUBUNIT: Heterooctamer of four A and four B subunits.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE GLUTACONATE COA-TRANSFERASE SUBUNIT B
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X81440; CAA57200.1; -.
 CC PIR; S51052; S51052.
 CC PDB; 1POI; 18-MAR-98.
 CC Transferase; 3D-structure.
 CC INIT MET 0 0
 CC ACT_SITE 53 53
 CC HELIX 8 17
 CC TURN 18 19
 CC TURN 22 23
 CC STRAND 25 27
 CC HELIX 32 42
 CC TURN 43 44
 CC TURN 46 47
 CC STRAND 49 52
 CC TURN 53 55
 CC STRAND 56 59
 CC TURN 68 69
 CC HELIX 71 74
 CC TURN 75 75
 CC STRAND 77 80
 CC HELIX 83 96
 CC TURN 97 97
 CC STRAND 101 105
 CC STRAND 109 110
 CC TURN 112 113
 CC STRAND 116 117
 CC STRAND 120 122
 CC STRAND 127 130
 CC TURN 135 136
 CC TURN 137 143
 CC STRAND 146 149
 CC TURN 154 156
 CC STRAND 157 157
 CC STRAND 166 166
 CC TURN 170 171
 CC TURN 173 174
 CC HELIX 175 178
 CC TURN 179 180
 CC TURN 183 184
 CC STRAND 187 192
 CC TURN 193 194
 CC STRAND 195 199
 CC TURN 201 203
 CC STRAND 206 211
 CC TURN 213 214
 CC HELIX 217 222
 CC TURN 223 223
 CC TURN 231 232
 CC STRAND 234 235
 CC HELIX 241 249
 CC TURN 250 251

FT TURN 253 254
SQ SEQUENCE 265 AA; 29035 MW; 1E7FF61B42162FB4 CRC64;
Query Match 39.4%; Score 43; DB 1; Length 265;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIEC 18
: ||| : ||| :
DB 62 VEVPSVGLRFRMAHCG 79
RESULT 23
SURE METMA STANDARD; PRT; 267 AA.
AC Q8PX31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acid phosphatase sure (EC 3.1.3.2).
GN SURE OR WM1391.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wlezer A., Baeumer S., Jacobi C.,
RA Bruggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -|- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
alcohol + phosphate.
CC -|- COFACTOR: Magnesium (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -|- SIMILARITY: Belongs to the sure acid phosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF013373; AAM31087.1; -.
DR HAMAP; MF 00060; -; 1.
DR InterPro; IPR002828; Sure.
DR Pfam; PF01975; Sure; 1.
DR ProDom; PD005378; Sure; 1.
DR TIGRFAMs; TIGR00087; sure; 1.
KW Hydrolase; Magnesium; Complete proteome.
FT ACT_SITE 133 133 POTENTIAL.
FT METAL 14 14 MAGNESIUM (BY SIMILARITY).
FT METAL 15 15 MAGNESIUM (BY SIMILARITY).
FT METAL 45 45 MAGNESIUM (BY SIMILARITY).
FT METAL 100 100 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 267 AA; 29097 MW; 709A83CC7854A771 CRC64;
Query Match 39.4%; Score 43; DB 1; Length 267;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 DDPRESSEDKRFESHIEC 17
: ||| : ||| :
DB 144 DDPDYQRFERFEAGIK 159
RESULT 24
RPBI_EUPOC STANDARD; PRT; 478 AA.
AC P28364;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase II largest subunit (EC 2.7.7.6)
DE (Fragment).
DB RPBI.
GN Euplates octocarinatus.
OS Eukaryota; Alveolata; Ciliophora; Spirotrichea; Hypotrichia;
OC Euplotida; Euplotidae; Euplates.
OX NCBI_TaxID=5937;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=(68)-VIII;
RX MEDLINE=93027138; PubMed=1408746;
RA Kaufmann J., Klein A.;
RT "Gene dosage as a possible major determinant for equal expression
levels of genes encoding RNA polymerase subunits in the hypotrichous
ciliate Euplates octocarinatus.";
RL Nucleic Acids Res. 20:4445-4450(1992).
CC -|- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
of DNA into RNA using the four ribonucleoside triphosphates as
substrates.
CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
(RNA)(N).
CC -|- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM TEN TO
TWELVE DIFFERENT POLYPEPTIDES. THIS POLYPEPTIDE IS THE LARGEST
COMPONENT OF RNA POLYMERASE II.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
found in eukaryotic nuclei: polymerase I for the ribosomal RNA
precursor, polymerase II for the mRNA precursor, and polymerase
III for 5S and tRNA genes.
CC -|- SIMILARITY: Belongs to the RNA polymerase beta' chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; X66452; CAA47068.1; -.
DR PIR; S13886; S13886.
DR InterPro; IPR000722; RNA_pol_A.
DR InterPro; IPR007080; RNA_pol_Rpbi_1.
DR InterPro; IPR008592; RNA_pol_A_N.
DR Pfam; PF04997; RNA_pol_Rpbi_1; 1.
DR Pfam; PF00623; RNA_pol_Rpbi_2; 1.
DR SMART; SM00663; RPOLA_N; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc; Repeat;
KW DNA-binding; Nuclear protein; Zinc-finger.
FT ZN_FING 68 84 C2H2-TYPE (POTENTIAL).
FT NON_TER 478 478
SQ SEQUENCE 478 AA; 54026 MW; 47A1ADD6ED98EB3D CRC64;
Query Match 39.4%; Score 43; DB 1; Length 478;
Best Local Similarity 47.4%; Pred. No. 43;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 2 DDPRESSEDKRFESHIEC 20
: ||| : ||| :
DB 186 DDPSTNDKRDLSASECLK 204
RESULT 25
CSID_ARATH STANDARD; PRT; 502 AA.
ID CSID_ARATH

FT TURN 253 254
SQ SEQUENCE 265 AA; 29035 MW; 1E7FF61B42162FB4 CRC64;
Query Match 39.4%; Score 43; DB 1; Length 265;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIEC 18
: ||| : ||| :
DB 62 VEVPSVGLRFRMAHCG 79
RESULT 23
SURE METMA STANDARD; PRT; 267 AA.
AC Q8PX31;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acid phosphatase sure (EC 3.1.3.2).
GN SURE OR WM1391.
OS Methanosarcina mazei (Methanosarcina frisia).
OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;
OC Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX MEDLINE=22120827; PubMed=12125824;
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
RA Martinez-Arias R., Henne A., Wlezer A., Baeumer S., Jacobi C.,
RA Bruggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC -|- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
alcohol + phosphate.
CC -|- COFACTOR: Magnesium (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -|- SIMILARITY: Belongs to the sure acid phosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF013373; AAM31087.1; -.
DR HAMAP; MF 00060; -; 1.
DR InterPro; IPR002828; Sure.
DR Pfam; PF01975; Sure; 1.
DR ProDom; PD005378; Sure; 1.
DR TIGRFAMs; TIGR00087; sure; 1.
KW Hydrolase; Magnesium; Complete proteome.
FT ACT_SITE 133 133 POTENTIAL.
FT METAL 14 14 MAGNESIUM (BY SIMILARITY).
FT METAL 15 15 MAGNESIUM (BY SIMILARITY).
FT METAL 45 45 MAGNESIUM (BY SIMILARITY).
FT METAL 100 100 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 267 AA; 29097 MW; 709A83CC7854A771 CRC64;
Query Match 39.4%; Score 43; DB 1; Length 267;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 DDPRESSEDKRFESHIEC 17
: ||| : ||| :
DB 144 DDPDYQRFERFEAGIK 159

AC Q9FG65; 065789;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Cytochrome P450 81D1 (EC 1.14.-.-)
 GN CYP81D1 OR CYP91A1 OR AT5G36220 OR T30G6.3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016721; PubMed=11130714;
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
 RA Yajima M., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
 RA Nakazaki N., Naruo K., Okumura S., Shinpo S., Takeuchi C., Wada T.,
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
 RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
 RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.A.,
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
 RA Volckaert G., Kambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
 RA Etian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
 RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
 RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
 RA van Staveren M., Dirkse W., Mooijman P., Klein Lankhorst R.,
 RA Weitzenecker T., Botche G., Rose M., Hauf J., Bernsner S., Hempel S.,
 RA Reldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
 RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis
 RA thaliana."
 RL Nature 408:823-826(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.B., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
 RA Arakawa T., Barth J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Heuan V.W., Iida K., Karnes M.,
 RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Ramey R., Vayberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RA "Empirical analysis of transcriptional activity in the Arabidopsis
 RA genome."
 RL Science 302:842-846(2003).
 RN [3]
 RP SEQUENCE OF 3-502 FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=98281573; PubMed=9620263;
 RA Mizutani M., Ward E., Ohta D.;
 RT "Cytochrome P450 superfamily in Arabidopsis thaliana. Isolation of
 RT cDNAs, differential expression, and RFLP mapping of multiple
 RT cytochromes P450."
 RL Plant Mol. Biol. 37:39-52(1998).
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB026661; BAB09361.1; -
 DR EMBL; D78606; BAA28538.1; -
 DR EMBL; AY093766; AAM10388.1; -
 DR HSSP; P14779; 1JFZ.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PRO0385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Transmembrane; Heme; Multigene family.
 FT TRANSMEM 6 26 POTENTIAL.
 FT METAL 440 440 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
 FT CONFLICT 121 121 T -> A (IN REF. 2).
 SQ SEQUENCE 502 AA; 56723 MW; 1A349A8DD76A41B0 CRC64;
 Query Match 39.4%; Score 43; DB 1; Length 502;
 Best Local Similarity 47.4%; Pred. No. 46;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 QY 2 DDPSEEDKRFESHIECKK 20
 DB 411 DDPDSFKPERKEEEAQQ 429
 RESULT 26
 DNAA FUSNN STANDARD; PRT; 607 AA.
 ID DNAA FUSNN STANDARD; PRT; 607 AA.
 AC Q8RH05;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Chapterone protein dnak (Heat shock protein 70) (Heat shock 70 kDa
 DE protein) (HSP70).
 DE protein)
 DN DNAA OR FN0116.
 OS Fusobacterium nucleatum (subsp. nucleatum).
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
 OC Fusobacterium.
 OC NCBI_TaxID=76856;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25586;
 RX MEDLINE=21886394; PubMed=11899109;
 RA Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
 RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
 RA Fongstein M., Kyripides N., Overbeek R.;
 RA "Genome sequence and analysis of the oral bacterium Fusobacterium
 RA nucleatum strain ATCC 25586."
 RT J. Bacteriol. 184:2005-2018(2002).
 CC -!- FUNCTION: Acts as a chaperone (By similarity).
 CC -!- INDUCTION: By stress conditions e.g. heat shock (By similarity).
 CC -!- SIMILARITY: Belongs to the heat shock protein 70 family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB010525; AAL94325.1; -
 DR HAMAP; MF 00332; -; 1.
 DR InterPro; IPR001023; Hsp70.
 DR Pfam; PF00012; HSP70; 1.
 DR PRINTS; PRO0301; HEATSHOCK70.

DR PRODM; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS00336; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;
FT MOD_RES 174 174 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 607 AA; 65440 MW; 307EL4BF42A35D1F CRC64;
Query Match 39.4%; Score 43; DB 1; Length 607;
Best Local Similarity 41.2%; Pred. No. 56;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY 3 DPRSEEDKRFESHIECR 19
DB 491 EAHAEEDKKFQELVEAR 507
RESULT 27
ACON HUMAN STANDARD; PRT; 780 AA.
AC Q99798; O75809;
DT 01-NOV-1997 (Rel. 35, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Aconitase hydratase, mitochondrial precursor (EC 4.2.1.3) (Citrate
DE hydro-lyase) (Aconitase).
GN ACO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Skeletal muscle;
RA Juang H.H., Chlou B.;
RT "Characterization of the human mitochondrial aconitase gene (ACO2).";
RL Gene 213:205-218(1998).
RP SEQUENCE FROM N.A.
RX MEDLINE=98296268; PubMed=9630632;
RA Mirel D.B., Marder K., Graziano J., Freyer G., Zhao Q., Mayeux R.,
RA Wilhelmsson K.C.;
RT "Characterization of the human mitochondrial aconitase gene (ACO2).";
RL Gene 213:205-218(1998).
RP SEQUENCE FROM N.A.
RC TISSUE-Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: Citrate = cis-aconitate + H(2)O.
CC -!- COFACTOR: Binds 1 4Fe-4S cluster per subunit. Binding of a 3Fe-4S
cluster leads to an inactive enzyme (By similarity).
CC -!- PATHWAY: Tricarboxylic acid cycle.
CC -!- SUBUNIT: Monomer (By similarity).

CC -!- SUBCELLULAR LOCATION: Mitochondrial (By similarity).
CC -!- SIMILARITY: Belongs to the aconitase/IFM isomerase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; U80040; AAC38416.1; -.
CC EMBL; U87939; AAC39921.1; -.
CC EMBL; U87926; AAC39921.1; JOINED.
CC EMBL; U87927; AAC39921.1; JOINED.
CC EMBL; U87928; AAC39921.1; JOINED.
CC EMBL; U87929; AAC39921.1; JOINED.
CC EMBL; U87930; AAC39921.1; JOINED.
CC EMBL; U87931; AAC39921.1; JOINED.
CC EMBL; U87932; AAC39921.1; JOINED.
CC EMBL; U87933; AAC39921.1; JOINED.
CC EMBL; U87934; AAC39921.1; JOINED.
CC EMBL; U87935; AAC39921.1; JOINED.
CC EMBL; U87936; AAC39921.1; JOINED.
CC EMBL; U87937; AAC39921.1; JOINED.
CC EMBL; U87938; AAC39921.1; JOINED.
CC EMBL; BC014092; AAH14092.1; -.
CC PIR; T52543; T52543.
CC HSP; P16276; 7ACN.
CC SWISS-2DPAGE; Q99798; HUMAN.
CC Genew; HGNC:118; ACO2.
CC GK; Q99798; -.
CC MIM; 100850; -.
CC GO; GO:0005739; C:mitochondrion; TAS.
CC GO; GO:0005506; F:iron ion binding; TAS.
CC GO; GO:0006101; P:citrate metabolism; TAS.
CC GO; GO:0006091; P:energy pathways; TAS.
CC InterPro; IPR000573; Aconitase_C.
CC InterPro; IPR006248; Aconitase_mito.
CC InterPro; IPR001030; Aconitase_N.
CC Pfam; PF00330; aconitase; 1.
CC Pfam; PF00694; aconitase_C; 1.
CC PRINTS; PR00415; ACONITASE.
CC ProDom; PD000511; Aconitase_N; 1.
CC TIGRFAMS; TIGR01340; aconitase_mito; 1.
CC PROSITE; PS00450; ACONITASE_1; 1.
CC PROSITE; PS01244; ACONITASE_2; 1.
KW Lyase; Tricarboxylic acid cycle; Iron-sulfur; Mitochondrion;
KW Transist peptide; 4Fe-4S; Pyroliadone carboxylic acid.
FT TRANSIT 1 27 MITOCHONDRION (By similarity).
FT CHAIN 28 780 ACONITASE HYDRATASE.
FT MOD_RES 28 28 PYRROLIDONE CARBOXYLIC ACID (By
FT METAL 385 385 IRON-SULFUR (4Fe-4S) (By similarity).
FT METAL 448 448 IRON-SULFUR (4Fe-4S) (By similarity).
FT METAL 451 451 IRON-SULFUR (4Fe-4S) (By similarity).
FT CONFLICT 35 35 S -> T (IN REF. 1).
FT CONFLICT 136 136 G -> D (IN REF. 1).
FT CONFLICT 159 159 A -> D (IN REF. 1).
FT CONFLICT 167 167 K -> S (IN REF. 1).
FT CONFLICT 242 242 S -> T (IN REF. 1).
FT CONFLICT 275 275 I -> M (IN REF. 1).
FT CONFLICT 517 517 T -> K (IN REF. 1).
FT CONFLICT 553 553 G -> R (IN REF. 1).
SQ SEQUENCE 780 AA; 85425 MW; 58C9FFBDDC63D5E CRC64;
Query Match 39.4%; Score 43; DB 1; Length 780;
Best Local Similarity 42.1%; Pred. No. 74;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECR 19
DB 74 LDDPASQEIERRGKSYLRLR 92

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RX MEDLINE=90307707; PubMed=2195026;
RA Winkelman J.C., Chang J.G., Tse W.T., Scarpa A.L., Marchesi V.T.,
RA Forget B.G.;
RT "Full-length sequence of the cDNA for human erythroid beta-spectrin.";
RL J. Biol. Chem. 265:11827-11832(1990).
RN [2]
RP SEQUENCE OF 1055-2137 FROM N.A. (ISOFORM 2).
RC TISSUE=Skeletal muscle;
RX MEDLINE=91056094; PubMed=2243099;
RA Winkelman J.C., Costa F., Linzie B.L., Forget B.G.;
RT "Beta spectrin in human skeletal muscle. Tissue-specific differential
RT processing of 3' beta spectrin pre-mRNA generates a beta spectrin
RT isoform with a unique carboxyl terminus.";
RL J. Biol. Chem. 265:20449-20454(1990).
RN [3]
RP SEQUENCE OF 2002-2137 FROM N.A.
RX MEDLINE=91332035; PubMed=1840591;
RA Gallagher P.G., Tse W.T., Costa F., Scarpa A., Boivin P., Delaunay J.,
RA Forget B.G.;
RT "A splice site mutation of the beta-spectrin gene causing exon
RT skipping in hereditary elliptocytosis associated with a truncated
RT beta-spectrin chain.";
RL J. Biol. Chem. 266:15154-15159(1991).
RN [4]
RP SEQUENCE OF 928-1755 FROM N.A.
RX MEDLINE=91007291; PubMed=1976574;
RA Yoon S.H., Kentros C.G., Prchal J.T.;
RT "Identification of an unusual deletion within homologous repeats of
RT human reticulocyte beta-spectrin and probable peptide polymorphism.";
RL Gene 91:297-302(1990).
RN [5]
RP SEQUENCE OF 1334-1432 AND 1903-2137 FROM N.A.
RX MEDLINE=88269838; PubMed=3390609;
RA Winkelman J.C., Leto T.L., Watkins P.C., Eddy R., Shows T.B., Forget B.G.;
RA Linenbach A.J., Sahr K.E., Kathuria N., Marchesi V.T., Forget B.G.;
RT "Molecular cloning of the cDNA for human erythrocyte beta-spectrin.";
RL Blood 72:328-334(1988).
RN [6]
RP SEQUENCE OF 1209-1482 FROM N.A.
RX MEDLINE=88041127; PubMed=3478706;
RA Prchal J.T., Morley B.J., Yoon S.H., Coetzer T.L., Palek J.,
RA Conboy J.G., Kan Y.W.;
RT "Isolation and characterization of cDNA clones for human erythrocyte
RT beta-spectrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7468-7472(1987).
RN [7]
RP DOMAINS.
RX MEDLINE=84295638; PubMed=6472478;
RA Speicher D.W., Marchesi V.T.;
RT "Erythrocyte spectrin is comprised of many homologous triple helical
RT segments.";
RL Nature 311:177-180(1984).
RN [8]
RP REVIEW ON VARIANTS.
RX MEDLINE=97001215; PubMed=8844207;
RA Maillet P., Alloisio N., Morle L., Delaunay J.;
RT "Spectrin mutations in hereditary elliptocytosis and hereditary
RT spherocytosis.";
RL Hum. Mutat. 8:97-107(1996).
RN [9]
RP VARIANT HE CAGLIARI GLY-2018.
RX MEDLINE=94043025; PubMed=8226774;
RA Sahr K.E., Coetzer T.L., Moy L.S., Derick L.H., Chishti A.H.,
RA Jarolim P., Lorenzo F., del Giudice E.M., Iolascon A., Gallanello R.,
RA Cao A., Delaunay J., Liu S.-C., Palek J.;
RT "Spectrin Cagliari: an Ala-->Gly substitution in helix 1 of beta
RT spectrin repeat 17 that severely disrupts the structure and self-
RT association of the erythrocyte spectrin heterodimer.";

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RX MEDLINE=90307707; PubMed=2195026;
RA Winkelman J.C., Chang J.G., Tse W.T., Scarpa A.L., Marchesi V.T.,
RA Forget B.G.;
RT "Full-length sequence of the cDNA for human erythroid beta-spectrin.";
RL J. Biol. Chem. 265:11827-11832(1990).
RN [2]
RP SEQUENCE OF 1055-2137 FROM N.A. (ISOFORM 2).
RC TISSUE=Skeletal muscle;
RX MEDLINE=91056094; PubMed=2243099;
RA Winkelman J.C., Costa F., Linzie B.L., Forget B.G.;
RT "Beta spectrin in human skeletal muscle. Tissue-specific differential
RT processing of 3' beta spectrin pre-mRNA generates a beta spectrin
RT isoform with a unique carboxyl terminus.";
RL J. Biol. Chem. 265:20449-20454(1990).
RN [3]
RP SEQUENCE OF 2002-2137 FROM N.A.
RX MEDLINE=91332035; PubMed=1840591;
RA Gallagher P.G., Tse W.T., Costa F., Scarpa A., Boivin P., Delaunay J.,
RA Forget B.G.;
RT "A splice site mutation of the beta-spectrin gene causing exon
RT skipping in hereditary elliptocytosis associated with a truncated
RT beta-spectrin chain.";
RL J. Biol. Chem. 266:15154-15159(1991).
RN [4]
RP SEQUENCE OF 928-1755 FROM N.A.
RX MEDLINE=91007291; PubMed=1976574;
RA Yoon S.H., Kentros C.G., Prchal J.T.;
RT "Identification of an unusual deletion within homologous repeats of
RT human reticulocyte beta-spectrin and probable peptide polymorphism.";
RL Gene 91:297-302(1990).
RN [5]
RP SEQUENCE OF 1334-1432 AND 1903-2137 FROM N.A.
RX MEDLINE=88269838; PubMed=3390609;
RA Winkelman J.C., Leto T.L., Watkins P.C., Eddy R., Shows T.B., Forget B.G.;
RA Linenbach A.J., Sahr K.E., Kathuria N., Marchesi V.T., Forget B.G.;
RT "Molecular cloning of the cDNA for human erythrocyte beta-spectrin.";
RL Blood 72:328-334(1988).
RN [6]
RP SEQUENCE OF 1209-1482 FROM N.A.
RX MEDLINE=88041127; PubMed=3478706;
RA Prchal J.T., Morley B.J., Yoon S.H., Coetzer T.L., Palek J.,
RA Conboy J.G., Kan Y.W.;
RT "Isolation and characterization of cDNA clones for human erythrocyte
RT beta-spectrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:7468-7472(1987).
RN [7]
RP DOMAINS.
RX MEDLINE=84295638; PubMed=6472478;
RA Speicher D.W., Marchesi V.T.;
RT "Erythrocyte spectrin is comprised of many homologous triple helical
RT segments.";
RL Nature 311:177-180(1984).
RN [8]
RP REVIEW ON VARIANTS.
RX MEDLINE=97001215; PubMed=8844207;
RA Maillet P., Alloisio N., Morle L., Delaunay J.;
RT "Spectrin mutations in hereditary elliptocytosis and hereditary
RT spherocytosis.";
RL Hum. Mutat. 8:97-107(1996).
RN [9]
RP VARIANT HE CAGLIARI GLY-2018.
RX MEDLINE=94043025; PubMed=8226774;
RA Sahr K.E., Coetzer T.L., Moy L.S., Derick L.H., Chishti A.H.,
RA Jarolim P., Lorenzo F., del Giudice E.M., Iolascon A., Gallanello R.,
RA Cao A., Delaunay J., Liu S.-C., Palek J.;
RT "Spectrin Cagliari: an Ala-->Gly substitution in helix 1 of beta
RT spectrin repeat 17 that severely disrupts the structure and self-
RT association of the erythrocyte spectrin heterodimer.";

RESULT 28

YQ36 CAEEL STANDARD; PRT; 963 AA.
AC Q09457;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative cuticle collagen C09G5.6.
GN C09G5.6
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Palmer S.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Nematode cuticles are composed largely of collagen-like
CC proteins. The cuticle functions both as an exoskeleton and as a
CC barrier to protect the worm from its environment (By similarity).
CC -!- SUBUNIT: Collagen polypeptide chains are complexed within the
CC cuticle by disulfide bonds and other types of covalent cross-links
CC (By similarity).
CC -!- SIMILARITY: Belongs to the cuticular collagen family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z46791; CAA86755.1; --
DR PIR: T19140; T19140.
DR WORMPep; C09G5.6; CE01486.
DR InterPro; IPR008161; C1g_helix.
DR InterPro; IPR002486; Col_cuticle_N.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR Pfam; PF01391; Collagen; 3.
DR ProDom; PD000007; C1g_helix; 2.
KW Hypothetical protein; Cuticle; Connective tissue; Repeat;
KW Multigene family; Collagen.
FT DOMAIN 392..423 TRIPLE-HELICAL REGION.
FT DOMAIN 441..503 TRIPLE-HELICAL REGION.
FT DOMAIN 506..567 TRIPLE-HELICAL REGION.
FT DOMAIN 663..666 POLY-PRO.
FT DOMAIN 685..688 POLY-PRO.
SQ SEQUENCE 963 AA; 107031 MW; AFF895A75909F66E CRC64;

Query Match 39.4%; Score 43; DB 1; Length 963;
Best Local Similarity 50.0%; Pred. No. 92;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PRSEEDKRFESHIECR 19
Db 777 PQEPRRYTEAPR 792

RESULT 29

SPCB HUMAN
ID SPCB HUMAN STANDARD; PRT; 2137 AA.
AC P11277; Q15510; Q15519;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Spectrin beta chain, erythrocyte (Beta-I spectrin).
GN SPTB OR SPB1.
OS Homo sapiens (Human).

RL J. Biol. Chem. 268:22656-22662(1993).
RN [10]
RP VARIANT HS KISSIMMEE ARG-202.
RX MEDLINE=93352802; PubMed=8102379;
RA Becker P.S., Tse W.T., Lux S.E., Forget B.G.;
RT "Beta spectrin KISSimmee: a spectrin variant associated with
RT autosomal dominant hereditary spherocytosis and defective binding to
RT protein 4.1";
RL J. Clin. Invest. 92:612-616(1993).
RN [11]
RP VARIANT HE PROVIDENCE PRO-2019.
RX MEDLINE=95190014; PubMed=7883966;
RA Gallagher P.G., Weed S.A., Tse W.T., Benoit L., Morrow J.S.,
RA Marchesi S.L., Mohandas N., Forget B.G.;
RT "Recurrent fetal hydrops fetalis associated with a nucleotide
RT substitution in the erythrocyte beta-spectrin gene";
RL J. Clin. Invest. 95:1174-1182(1995).
RN [12]
RP VARIANTS HE VAL-2023 AND ARG-2024.
RX MEDLINE=94289716; PubMed=8018926;
RA Parquet N., Devaux I., Boulanger L., Galland C., Boivin P.,
RA Lecomte M.-C., Dhery D., Garbarz M.;
RT "Identification of three novel spectrin alpha I/74 mutations in
RT hereditary elliptocytosis: further support for a triple-stranded
RT folding unit model of the spectrin heterodimer contact site";
RL Blood 84:303-308(1994).
RN [13]
RP VARIANT HE PRO-2053.
RX MEDLINE=90369011; PubMed=1375598;
RA Tse W.T., Lecomte M.-C., Costa F.F., Garbarz M., Feo C., Boivin P.,
RA Dhery D., Forget B.G.;
RT "Point mutation in the beta-spectrin gene associated with alpha I/74
RT hereditary elliptocytosis. Implications for the mechanism of spectrin
RT dimer self-association";
RL J. Clin. Invest. 86:909-916(1990).
CC -!- FUNCTION: Spectrin is the major constituent of the cytoskeletal
CC network underlying the erythrocyte plasma membrane. It associates
CC with band 4.1 and actin to form the cytoskeletal superstructure of
CC the erythrocyte plasma membrane.
CC -!- SUBUNIT: Composed of nonhomologous chains, alpha and beta, which
CC aggregate to form dimers, tetramers, and higher polymers.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=P11277-1; Sequences=Displayed;
CC Name=2; Synonyms=Muscle-specific;
CC IsoId=P11277-2; Sequences=VSP_000719;
CC Name=3;
CC IsoId=P11277-3; Sequences=VSP_007242;
CC -!- DISEASE: Defects in SPTB are a cause of rhesus-unlinked hereditary
CC elliptocytosis (HE) [MIM:182870, 130600, 166900]; also called
CC elliptocytosis 2 (EL2). Hereditary elliptocytosis (also known as
CC hereditary ovalocytosis) is a genetically heterogeneous, autosomal
CC dominant hemologic disorder. It is characterized by variable
CC hemolytic anemia and elliptical or oval red cell shape. Several
CC abnormalities in the membrane skeleton have been identified,
CC including a number that have been localized to spectrin.
CC Elliptocytosis due to defects in SPTB can also be referred to as
CC elliptocytosis 3 (EL3).
CC -!- DISEASE: Defects in SPTB are the cause of spherocytosis type I
CC (SPH1) [MIM:182870], a disorder characterized by severe hemolytic
CC anemia. Inheritance is autosomal dominant.
CC -!- MISCELLANEOUS: This complex is anchored to the cytoplasmic face of
CC the plasma membrane via another protein, ankyrin, which binds to
CC beta-spectrin and mediates the binding of the whole complex to a
CC transmembrane protein band 3. The interaction of erythrocyte
CC spectrin with other proteins through specific binding domains lead
CC to the formation of an extensive subplasmalemmal meshwork which is
CC thought to be responsible for the maintenance of the biconcave
CC shape of human erythrocytes, for the regulation of plasma membrane
CC components and for the maintenance of the lipid asymmetry of the
CC plasma membrane.
CC -!- SIMILARITY: Belongs to the spectrin family.

CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 17 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J05500; AAA60578.1; --
CC EMBL: J05500; AAA60579.1; --
CC EMBL: M37884; AAA63259.1; --
CC EMBL: M37885; AAA60571.1; --
CC EMBL: M57948; --; NOT_ANNOTATED_CDS.
CC EMBL: X59510; CAA42097.1; --
CC EMBL: X59511; CAA42098.1; --
CC EMBL: M18054; AAA60572.1; --
CC PIR: A37064; SJHJB.
CC HSP: O01082; 1BKR.
CC SWISS-2DPAGE: P11277; HUMAN.
CC Genew: HGNC:11274; SPTB.
CC MIM: 182870; --
CC MIM: 130600; --
CC MIM: 166900; --
CC GO: GO:0008091; C-spectrin; TAS.
CC GO: GO:0003779; F-actin binding; TAS.
CC InterPro: IPR001589; Actbind actinin.
CC InterPro: IPR001715; Calponin-like.
CC InterPro: IPR002017; Spectrin.
CC Pfam: PF00307; CH; 2.
CC Pfam: PF00435; spectrin; 17.
CC SMART: SM00033; CH; 2.
CC SMART: SM00150; SPEC; 17.
CC PROSITE: PS00019; ACTININ_1; 1.
CC PROSITE: PS00020; ACTININ_2; 1.
CC PROSITE: PS50021; CH; 2.
CC Cytoskeleton; Membrane; Erythrocyte; Repeat; Actin-binding;
CC Disease mutation; Hereditary hemolytic anemia; Elliptocytosis;
CC Alternative splicing; Actin capping.
CC DOMAIN 1 275 ACTIN-BINDING.
CC FT DOMAIN 54 158 CH 1.
CC FT DOMAIN 173 275 CH 2.
CC FT REPEAT 276 384 SPECTRIN 1.
CC FT REPEAT 385 498 SPECTRIN 2.
CC -----
CC Query Match 39.4%; Score 43; DB 1; Length 2137;
CC Best Local Similarity 43.8%; Pred. No. 2.2e+02;
CC Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
CC -----
CC Qy 4 PRSEEDKRFESHIECR 19
CC ||| |:::| |
CC Db 920 PRSEVKVQYQDHLNTR 935
CC -----
CC RESULT 30
CC SFAS_CHLEU STANDARD; PRT; 269 AA.
CC ID_Q39490;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DB SF-assemblin.
CC OS Chlamydomonas eugametos.
CC OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
CC OC Chlamydomonadaceae; Chlamydomonas.
CC OX NCBI_TaxID=3053;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA MEDLINE=97167588; PubMed=9015206;
CC RA Lechtreck K.-F., Silflow C.;
CC RT "SF-assemblin in Chlamydomonas: sequence conservation and localization

```
RT during the cell cycle."
RL Cell Motil. Cytoskeleton 36:190-201(1997).
CC -!- FUNCTION: Major component of the striated microtubule-associated
CC fibers (SMAFs; system-1-fibers).
CC -!- DOMAIN: Consists of a small non-helical N-terminal domain and a
CC rod domain with a 29 residue repeat pattern based on four heptads
CC followed by a skip residue. This alpha-helical protein is
CC characterized by the ability to form a special segmented coiled
CC coil and to assemble into striated fibers of 2 nm protofilaments
CC (By similarity).
CC -!- SIMILARITY: TO OTHER GREEN ALGAE SF-ASSEMBLIN AND GIARDIA
CC LAMBLLA BETA-GIARDIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; U56981; AAB03788.1; -.
CC InterPro: IPR008374; SF assemblin.
CC PRINTS; PR01799; SPASSEMBLIN.
CC KW Cytoskeleton; Microtubule; Coiled coil.
CC FT DOMAIN 1 24 NONHELICAL REGION.
CC FT DOMAIN 25 269 ROD.
CC FT DOMAIN 98 144 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 269 AA; 30444 MW; 38CE792501C307D0 CRC64;
CC
CC Query Match 38.5%; Score 42; DB 1; Length 269;
CC Best Local Similarity 53.8%; Pred.No. 33;
CC Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
CC
CC Qy 5 RSEEDKRFESHIE 17
CC Db 79 RAESDKQLQSHFE 91
```

Search completed: May 11, 2004, 13:37:58
Job time : 8.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:28:26 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-09-171-432A-38
Perfect score: 109
Sequence: 1 VDDPSHEDKRFESHIECRK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : SPTREMBL_25.*

- 1: sp archaea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp mhc.*
- 8: sp organelle.*
- 9: sp phage.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp rvirus.*
- 16: sp bacteriaph.*
- 17: sp archaea.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	109	100.0	56	12	Q67819 hepatitis a
2	109	100.0	56	12	Q67818 hepatitis a
3	109	100.0	56	12	Q67820 hepatitis a
4	109	100.0	56	12	Q39865 hepatitis a
5	109	100.0	56	12	Q7ICYS hepatitis a
6	109	100.0	56	12	Q8JMR7 hepatitis a
7	109	100.0	56	12	Q8JMS2 hepatitis a
8	109	100.0	56	12	Q8JN35 hepatitis a
9	109	100.0	56	12	Q8C97 hepatitis a
10	109	100.0	56	12	Q8C96 hepatitis a
11	109	100.0	56	12	Q9GCT1 hepatitis a
12	109	100.0	56	12	Q9GCT2 hepatitis a
13	109	100.0	56	12	Q92941 hepatitis a
14	109	100.0	56	12	Q858K3 hepatitis a
15	109	100.0	56	12	Q8V4L9 hepatitis a
16	109	100.0	56	12	Q8V4L6 hepatitis a
17	109	100.0	56	12	Q8V4J9 hepatitis a
18	109	100.0	56	12	Q8V4J5 hepatitis a
19	109	100.0	56	12	Q8V4N3 hepatitis a
20	109	100.0	56	12	Q8V4M4 hepatitis a
21	109	100.0	56	12	Q8V4N6 hepatitis a
22	109	100.0	56	12	Q8V4K3 hepatitis a
23	109	100.0	56	12	Q8V4M2 hepatitis a
24	109	100.0	56	12	Q8V4N0 hepatitis a
25	109	100.0	56	12	Q8V4K8 hepatitis a
26	109	100.0	56	12	Q8V4K0 hepatitis a
27	109	100.0	56	12	Q8V4M1 hepatitis a
28	109	100.0	56	12	Q8V4K5 hepatitis a
29	109	100.0	56	12	Q8B8M1 hepatitis a
30	109	100.0	56	12	Q8B8L7 hepatitis a
31	109	100.0	56	12	Q8V3Z7 hepatitis a
32	109	100.0	56	12	Q8V403 hepatitis a
33	109	100.0	56	12	Q8V402 hepatitis a
34	109	100.0	56	12	Q8V3Z6 hepatitis a
35	109	100.0	56	12	Q8V411 hepatitis a
36	109	100.0	56	12	Q8V3Z9 hepatitis a
37	109	100.0	56	12	Q8V408 hepatitis a
38	109	100.0	56	12	Q8V404 hepatitis a
39	109	100.0	56	12	Q8V410 hepatitis a
40	109	100.0	56	12	Q8V412 hepatitis a
41	109	100.0	56	12	Q8V3Z8 hepatitis a
42	109	100.0	56	12	Q8V409 hepatitis a
43	109	100.0	56	12	Q8V3Z5 hepatitis a
44	109	100.0	56	12	Q8V407 hepatitis a
45	109	100.0	56	12	Q8V400 hepatitis a
46	109	100.0	56	12	Q8V406 hepatitis a
47	109	100.0	56	12	Q8V405 hepatitis a
48	109	100.0	56	12	Q8V401 hepatitis a
49	109	100.0	56	12	Q8V403 hepatitis a
50	109	100.0	56	12	Q8V404 hepatitis a
51	109	100.0	56	12	Q8V410 hepatitis a
52	109	100.0	56	12	Q8V412 hepatitis a
53	109	100.0	56	12	Q8V3Z8 hepatitis a
54	109	100.0	56	12	Q8V409 hepatitis a
55	109	100.0	56	12	Q8V3Z5 hepatitis a
56	109	100.0	56	12	Q8V407 hepatitis a
57	109	100.0	56	12	Q8V400 hepatitis a
58	109	100.0	56	12	Q8V406 hepatitis a
59	109	100.0	56	12	Q8V405 hepatitis a
60	109	100.0	56	12	Q8V401 hepatitis a
61	109	100.0	56	12	Q8V403 hepatitis a
62	109	100.0	56	12	Q8V404 hepatitis a
63	109	100.0	56	12	Q8V410 hepatitis a
64	109	100.0	56	12	Q8V412 hepatitis a
65	109	100.0	56	12	Q8V3Z8 hepatitis a
66	109	100.0	56	12	Q8V409 hepatitis a
67	109	100.0	56	12	Q8V3Z5 hepatitis a
68	109	100.0	56	12	Q8V407 hepatitis a
69	109	100.0	56	12	Q8V400 hepatitis a
70	109	100.0	56	12	Q8V406 hepatitis a
71	109	100.0	56	12	Q8V405 hepatitis a
72	109	100.0	56	12	Q8V401 hepatitis a
73	109	100.0	56	12	Q8V403 hepatitis a
74	109	100.0	56	12	Q8V404 hepatitis a
75	109	100.0	56	12	Q8V410 hepatitis a
76	109	100.0	56	12	Q8V412 hepatitis a
77	109	100.0	56	12	Q8V3Z8 hepatitis a
78	109	100.0	56	12	Q8V409 hepatitis a
79	109	100.0	56	12	Q8V3Z5 hepatitis a
80	109	100.0	56	12	Q8V407 hepatitis a
81	109	100.0	56	12	Q8V400 hepatitis a
82	109	100.0	56	12	Q8V406 hepatitis a
83	109	100.0	56	12	Q8V405 hepatitis a
84	109	100.0	56	12	Q8V401 hepatitis a
85	109	100.0	56	12	Q8V403 hepatitis a
86	109	100.0	56	12	Q8V404 hepatitis a
87	109	100.0	56	12	Q8V410 hepatitis a
88	109	100.0	56	12	Q8V412 hepatitis a
89	109	100.0	56	12	Q8V3Z8 hepatitis a
90	109	100.0	56	12	Q8V409 hepatitis a
91	109	100.0	56	12	Q8V3Z5 hepatitis a
92	109	100.0	56	12	Q8V407 hepatitis a
93	109	100.0	56	12	Q8V400 hepatitis a
94	109	100.0	56	12	Q8V406 hepatitis a
95	109	100.0	56	12	Q8V405 hepatitis a
96	109	100.0	56	12	Q8V401 hepatitis a
97	109	100.0	56	12	Q8V403 hepatitis a
98	109	100.0	56	12	Q8V404 hepatitis a
99	109	100.0	56	12	Q8V410 hepatitis a
100	109	100.0	56	12	Q8V412 hepatitis a

90 106 97.2 56 12 Q8JYK3 hepatitis a
 91 106 97.2 56 12 Q9IPD0 hepatitis a
 92 106 97.2 56 12 Q9IPB3 hepatitis a
 93 106 97.2 56 12 Q9IPB2 hepatitis a
 94 106 97.2 56 12 Q99U7 hepatitis a
 95 106 97.2 56 12 Q99H1 hepatitis a
 96 106 97.2 56 12 Q99H1 hepatitis a
 97 106 97.2 56 12 Q99H1 hepatitis a
 98 106 97.2 56 12 Q99H1 hepatitis a
 99 106 97.2 56 12 Q99H1 hepatitis a
 100 106 97.2 56 12 Q99U6 hepatitis a

ALIGNMENTS

RESULT 1
 Q67822 PRELIMINARY; PRT; 56 AA.
 ID Q67822
 AC Q67822
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISR-70;
 RA Karetvi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
 RA Mendelson E.;
 RT "Genetic classification of hepatitis A virus strains isolated in
 RT Israel, based on their VP1/2A nucleotide sequence."
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z77247; CAB01040.1; -;
 DR PIR; Q0427; Q0427.
 DR PIR; Q0428; Q0428.
 DR PIR; Q0431; Q0431.
 DR InterPro; IPR000886; ER target_S.
 DR PROSITE; PS00014; ER TARGET; 1.
 FT NON TER 1
 FT CHAIN 1 >28 CAPSID PROTEIN VP1.
 FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
 FT NON TER 56
 SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
 Best Local Similarity 100.0%; Pred. No. 5.8e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
 DB 16 VDDPRSEEDKRFESHIECKR 35
 RESULT 2
 Q39867 PRELIMINARY; PRT; 56 AA.
 ID Q39867
 AC Q39867
 DT 01-JAN-1998 (Tremblrel. 05, Created)
 DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JVR;
 RX MEDLINE=97247817; PubMed=9093940;

RA Taylor M.B.;
 RT "Molecular epidemiology of South African strains of hepatitis A virus:
 RT 1982-1996.";
 RL J. Med. Virol. 51:273-279(1997).
 DR EMBL; U68692; AAB53588.1; -;
 DR PIR; Q0427; Q0427.
 DR PIR; Q0428; Q0428.
 DR PIR; Q0431; Q0431.
 DR InterPro; IPR000886; ER target_S.
 DR PROSITE; PS00014; ER TARGET; 1.
 FT NON TER 1
 FT NON TER 56
 SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
 Best Local Similarity 100.0%; Pred. No. 5.8e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
 DB 16 VDDPRSEEDKRFESHIECKR 35
 RESULT 3
 Q918Q3 PRELIMINARY; PRT; 56 AA.
 ID Q918Q3
 AC Q918Q3
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RJ-005;
 RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
 RT "Characterization of hepatitis A virus isolates from subgenotypes IA
 RT and IB in Rio de Janeiro, Brazil.";
 RL J. Med. Virol. 0:0-0(2001).
 DR EMBL; AF410381; AAL0177.1; -;
 DR PIR; Q0427; Q0427.
 DR PIR; Q0428; Q0428.
 DR PIR; Q0431; Q0431.
 DR InterPro; IPR000886; ER target_S.
 DR PROSITE; PS00014; ER TARGET; 1.
 FT NON TER 1
 FT NON TER 56
 SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
 Best Local Similarity 100.0%; Pred. No. 5.8e-10;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
 DB 16 VDDPRSEEDKRFESHIECKR 35
 RESULT 4
 Q91P95 PRELIMINARY; PRT; 56 AA.
 ID Q91P95
 AC Q91P95
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 CC Hepatovirus.
 OX NCBI_TaxID=12092;

RN SEQUENCE FROM N.A.
RP STRAIN=A7;
RC FUJIWARA K.;
RX "hepatitis A virus VP1/2A junction."
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RT EMBL; AB046915; BAB08072.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 5
Q91800 PRELIMINARY; PRT; 56 AA.
ID Q91800;
AC Q91800;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RJ-055;
RA de Paula V.S., Baptista M.L., Lampe B., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410384; AALI0180.1; -
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6598 MW; E838C51846AEF4B4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 6
Q91800 PRELIMINARY; PRT; 56 AA.
ID Q91800;
AC Q91800;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;

[1] SEQUENCE FROM N.A.
RP STRAIN=96001190;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68690; AAB53586.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 7
Q67821 PRELIMINARY; PRT; 56 AA.
ID Q67821;
AC Q67821;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-15;
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
Israel, based on their VP1/2A nucleotide sequence."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77246; CAB01039.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 8
Q91804 PRELIMINARY; PRT; 56 AA.
ID Q91804
AC Q91804;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)


```
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Polypeptide (Fragment)
Hepatitis A virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
STRAIN=RU-004;
de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
"Characterization of Hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil.";
J. Med. Virol. 0:0-0(2001).
EMBL; AF410380; AAL10176.1; -.
PIR; PQ0427; PQ0427.
PIR; PQ0428; PQ0428.
PIR; PQ0431; PQ0431.
InterPro; IPR000886; ER target S.
PROSITE; PS00014; ER_TARGET; 1.
NON_TER 1
NON_TER 56
SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 9
Q91PA8 PRELIMINARY; PRT; 56 AA.
AC Q91PA8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
DE Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
STRAIN=A40;
Fujiwara K.;
"hepatitis A virus VP1/2A junction.";
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AB046902; BAB08059.1; -.
PIR; PQ0427; PQ0427.
PIR; PQ0428; PQ0428.
PIR; PQ0431; PQ0431.
InterPro; IPR000886; ER target S.
PROSITE; PS00014; ER_TARGET; 1.
NON_TER 1
NON_TER 56
SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 10
Q67823 PRELIMINARY; PRT; 56 AA.
ID Q67823
```

```
Q67823;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
DE Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
STRAIN=ISR-79;
Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
Mendelson E.;
"Genetic classification of hepatitis A virus strains isolated in
Israel, based on their VP1/2A nucleotide sequence.";
Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
EMBL; Z77248; CAB01041.1; -.
PIR; PQ0427; PQ0427.
PIR; PQ0428; PQ0428.
PIR; PQ0431; PQ0431.
InterPro; IPR000886; ER target S.
PROSITE; PS00014; ER_TARGET; 1.
NON_TER 1
NON_TER 56
CHAIN 29 >56 CAPSID PROTEIN VP1.
NONSTRUCTURAL PROTEIN NS2A.
PIR; PQ0431; PQ0431.
NON_TER 56
SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 11
Q39874 PRELIMINARY; PRT; 56 AA.
ID Q39874;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment)
DE Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
[1]
SEQUENCE FROM N.A.
STRAIN=314274;
MEDLINE=97247817; PubMed=9093940;
Taylor M.B.;
"Molecular epidemiology of South African strains of hepatitis A virus:
1982-1996.";
J. Med. Virol. 51:273-279(1997).
EMBL; U68699; AAB53595.1; -.
PIR; PQ0427; PQ0427.
PIR; PQ0428; PQ0428.
PIR; PQ0431; PQ0431.
InterPro; IPR000886; ER target S.
PROSITE; PS00014; ER_TARGET; 1.
NON_TER 1
NON_TER 56
SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
```

Db 16 VDDPRSEEDKRFESHIECRK 35
|||||

RESULT 12
Q918P9 PRELIMINARY; PRT; 56 AA.
ID Q918P9
AC Q918P9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RJ-504;
RA de Paula V.S.; Baptista M.L.; Lampe E.; Niel C.; Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
RT and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410385; AAL10181.1; -.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6598 MW; E938C51846AEF4B4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

RESULT 13
Q918P9 PRELIMINARY; PRT; 56 AA.
ID Q918P9
AC Q918P9
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RJ-504;
RA de Paula V.S.; Baptista M.L.; Lampe E.; Niel C.; Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
RT and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68694; AAB53590.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

RESULT 14
Q918P9 PRELIMINARY; PRT; 56 AA.
ID Q918P9
AC Q918P9
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=503712;
RA MEDLINE=97247817; PubMed=9093940;
RT Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RT 1982-1996."
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68693; AAB53589.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||

RESULT 15
Q918Q1 PRELIMINARY; PRT; 56 AA.
ID Q918Q1
AC Q918Q1
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RJ-NSG2;
RA de Paula V.S.; Baptista M.L.; Lampe E.; Niel C.; Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
RT and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 0:0-0(2001).
DR EMBL; AF410383; AAL10179.1; -.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 698E074846AEF4B7 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 16

O19873 ID O39873 PRELIMINARY; PRT; 56 AA.
AC O39873;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2333;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68698; AB53594.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 17

O67819 ID O67819 PRELIMINARY; PRT; 56 AA.
AC O67819;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-10;
RA Karetnyi Y.V.; Shulman L.M.; Manor J.; Leitner L.; Shehab S.;
RA Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z77244; CAB01037.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 18

O67818 ID O67818 PRELIMINARY; PRT; 56 AA.
AC O67818;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-05;
RA Karetnyi Y.V.; Shulman L.M.; Manor J.; Leitner L.; Shehab S.;
RA Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z77243; CAB01036.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON TER 1
FT NON TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 19

O67820 ID O67820 PRELIMINARY; PRT; 56 AA.
AC O67820;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-12;
RA Karetnyi Y.V.; Shulman L.M.; Manor J.; Leitner L.; Shehab S.;
RA Mendelson E.;

FT NON TER 1 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 18

O67818 ID O67818 PRELIMINARY; PRT; 56 AA.
AC O67818;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-05;
RA Karetnyi Y.V.; Shulman L.M.; Manor J.; Leitner L.; Shehab S.;
RA Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z77243; CAB01036.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||
Db 16 VDDPRSEEDKRFESHIECRK 35

RESULT 19

O67820 ID O67820 PRELIMINARY; PRT; 56 AA.
AC O67820;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-12;
RA Karetnyi Y.V.; Shulman L.M.; Manor J.; Leitner L.; Shehab S.;
RA Mendelson E.;

RT "Genetic classification of hepatitis A virus strains isolated in
RT Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77245; CAB01038.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 -1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438CS1846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 20
C39865 PRELIMINARY; PRT; 56 AA.
AC C39865;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=412991;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RT 1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68689; AAB53585.1; -.
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 -1
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438CS1846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 21
Q7TCY5 PRELIMINARY; PRT; 56 AA.
AC Q7TCY5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg9.6;
RA Munne M.S., Vladimirov S.N., Gonzalez J.E.;
RT "Molecular characterization of hepatitis A virus from Argentina.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY314885; AAF79464.1; -.
DR Polyprotein.
FT NON_TER 1 1
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438CS1846AEF4A4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 5.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||
DB 16 VDDPRSEEDKRFESHIECKR 35

RESULT 22
Q8JMR7 PRELIMINARY; PRT; 59 AA.
AC Q8JMR7;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A'dam085;
RA ter Bruggen J.P., Tjon G.M.S., Koek A.G., Wijkman C.J.,
RA van Lieshout I.J., Leenders A.C.A.P., Bruisten S.M.;
RT "Direct evidence of hepatitis A virus involvement in meningitis.";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY101271; AAM50074.1; -.
FT NON_TER 1 1
FT NON_TER 59 59
SQ SEQUENCE 59 AA; 6792 MW; FF358FBCD4F89D7 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 59;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
|||
DB 23 VDDPRSEEDKRFESHIECKR 42

RESULT 23
Q8JMS2 PRELIMINARY; PRT; 59 AA.
AC Q8JMS2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Brab008;
RA ter Bruggen J.P., Tjon G.M.S., Koek A.G., Wijkman C.J.,
RA van Lieshout I.J., Leenders A.C.A.P., Bruisten S.M.;

RT "Direct evidence of hepatitis A virus involvement in meningitis.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY101266; AAM50069.1; -
FT NON_TER 1
FT NON_TER 59
SQ SEQUENCE 59 AA; 6820 MW; ED8EE8FBCD4F99D7 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 59;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 23 VDDPRSEEDKRFESHIECRK 42

RESULT 24
Q8JUN35 PRELIMINARY; PRT; 71 AA.
AC Q8JUN35;
DT 01-OCT-2002 (TremBLrel. 22, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAF300;
RA De Paula V.S., Saback F.L., Gaspar A.M.C., Niel C.;
RT "Mixed Infection of a Young Female Child Care Provider With Hepatitis A Virus Isolates From Subgenotypes IA and IB Revealed by Heteroduplex Mobility Assay.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY083511; AAM03320.1; -
FT NON_TER 1
FT NON_TER 71
SQ SEQUENCE 71 AA; 8322 MW; 05D9E23399F3E446 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 71;
Best Local Similarity 100.0%; Pred. No. 7.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 32 VDDPRSEEDKRFESHIECRK 51

RESULT 25
Q8BC97 PRELIMINARY; PRT; 72 AA.
AC Q8BC97;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAF163;
RA Villar L.M., Lampe E., Meyer A., Gaspar A.M.C.;
RT "Genetic variability of hepatitis A virus strains isolated during an outbreak in a public school in Rio de Janeiro, Brazil.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF538726; AAN15169.1; -
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8425 MW; 06C6DAE2309AF3E4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 7.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 32 VDDPRSEEDKRFESHIECRK 51

RESULT 26
Q8BC96 PRELIMINARY; PRT; 72 AA.
AC Q8BC96;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAF300;
RA Villar L.M., Lampe E., Meyer A., Gaspar A.M.C.;
RT "Genetic variability of hepatitis A virus strains isolated during an outbreak in a public school in Rio de Janeiro, Brazil.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF538727; AAN15169.1; -
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8425 MW; 06C6DAE2309AF3E4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 7.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 32 VDDPRSEEDKRFESHIECRK 51

RESULT 27
Q9QCT0 PRELIMINARY; PRT; 97 AA.
AC Q9QCT0;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DE Capsid protein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIB4022;
RA Nainan O.V., Hutin Y.J., Margolis H.S., Alter M.J.;
RT "A Multistate, Foodborne Outbreak of Hepatitis A.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199046; AAF08937.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 97
SQ SEQUENCE 97 AA; 11282 MW; 5EC8BF22812A62B9 CRC64;

FT NON_TER 72 72
SQ SEQUENCE 72 AA; 8435 MW; 06C5D9E23399F3E4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 7.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 32 VDDPRSEEDKRFESHIECRK 51

RESULT 26
Q8BC96 PRELIMINARY; PRT; 72 AA.
AC Q8BC96;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAF300;
RA Villar L.M., Lampe E., Meyer A., Gaspar A.M.C.;
RT "Genetic variability of hepatitis A virus strains isolated during an outbreak in a public school in Rio de Janeiro, Brazil.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF538727; AAN15169.1; -
FT NON_TER 1
FT NON_TER 72
SQ SEQUENCE 72 AA; 8425 MW; 06C6DAE2309AF3E4 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 72;
Best Local Similarity 100.0%; Pred. No. 7.5e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
DB 32 VDDPRSEEDKRFESHIECRK 51

RESULT 27
Q9QCT0 PRELIMINARY; PRT; 97 AA.
AC Q9QCT0;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DE Capsid protein (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIB4022;
RA Nainan O.V., Hutin Y.J., Margolis H.S., Alter M.J.;
RT "A Multistate, Foodborne Outbreak of Hepatitis A.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199046; AAF08937.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 97
SQ SEQUENCE 97 AA; 11282 MW; 5EC8BF22812A62B9 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
Db 47 VDDPRSEEDKRFESHIECKR 66

RESULT 28
Q9QCT1 PRELIMINARY; PRT; 97 AA.
ID Q9QCT1
AC Q9QCT1;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB35;
RA Nainan O.V., Hutin Y.J., Margolis H.S., Alter M.J.;
RT "A Multistate, Foodborne Outbreak of Hepatitis A";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199045; AAF08935.1; -;
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 97
SQ SEQUENCE 97 AA; 11264 MW; 57C216228E259246 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
Db 47 VDDPRSEEDKRFESHIECKR 66

RESULT 29
Q9QCT2 PRELIMINARY; PRT; 97 AA.
ID Q9QCT2
AC Q9QCT2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MB35;
RA Nainan O.V., Hutin Y.J., Margolis H.S., Alter M.J.;
RT "A Multistate, Foodborne Outbreak of Hepatitis A";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF199044; AAF08935.1; -;
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 97
SQ SEQUENCE 97 AA; 11282 MW; 5ECB8F22812A62B9 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 97;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
Db 47 VDDPRSEEDKRFESHIECKR 66

RESULT 30
Q92941 PRELIMINARY; PRT; 116 AA.
ID Q92941
AC Q92941;
DT 01-NOV-1998 (TREMELrel. 08, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-1;
RX MEDLINE=97442186; PubMed=9298735;
RA Grinde B., Stene-Johansen K., Sharma B., Hoel T., Jensenius M.,
RA Skaug K.;
RT "Characterisation of an epidemic of hepatitis A virus involving
RT intravenous drug abusers--infection by needle sharing?";
RL J. Med. Virol. 53:69-75 (1997).
DR EMBL; AF050223; AAC05516.1; -;
DR EMBL; AF050231; AAC05524.1; -;
DR EMBL; AF050229; AAC05522.1; -;
DR EMBL; AF050230; AAC05523.1; -;
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
KW Polyprotein.
FT NON_TER 1
FT NON_TER 116
SQ SEQUENCE 116 AA; 13479 MW; 4726C6A1E7EBA64 CRC64;

Query Match 100.0%; Score 109; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECKR 20
Db 52 VDDPRSEEDKRFESHIECKR 71

Search completed: May 11, 2004, 13:40:21
Job time : 36 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:34:17 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-09-171-432A-38

Perfect score: 109

Sequence: 1 VDDPRSEDKRPFESIECRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/2/1aa/6CTUS_COMB.pep:*
- 6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	1091	6	Patent No. 5516630
2	109	100.0	2227	3	Sequence 2, Appli
3	109	100.0	2227	3	Sequence 4, Appli
4	109	100.0	2227	3	Sequence 6, Appli
5	109	100.0	2227	3	Sequence 2, Appli
6	109	100.0	2227	3	Sequence 4, Appli
7	109	100.0	2227	3	Sequence 2, Appli
8	109	100.0	2227	4	Sequence 2, Appli
9	109	100.0	2227	4	Sequence 4, Appli
10	109	100.0	2227	4	Sequence 6, Appli
11	109	100.0	2227	4	Sequence 2, Appli
12	109	100.0	2227	4	Sequence 4, Appli
13	109	100.0	2227	4	Sequence 2, Appli
14	109	97.2	2227	4	Sequence 6, Appli
15	81	74.3	839	1	Sequence 12, Appli
16	48.5	44.5	112	4	Sequence 3240, Ap
17	46	42.2	567	4	Sequence 21426, A
18	44.5	40.8	1334	2	Sequence 2, Appli
19	44.5	40.8	1334	3	Sequence 2, Appli
20	44	40.4	185	4	Sequence 30750, A
21	44	40.4	496	4	Sequence 31407, A
22	43.5	39.9	818	4	Sequence 5208, Ap
23	43	39.4	780	2	Sequence 2, Appli
24	42	38.5	180	6	Patent No. 5273901
25	42	38.5	180	6	Patent No. 5482709
26	42	38.5	560	2	Sequence 5, Appli
27	42	38.5	786	4	Sequence 23, Appli

28	42	38.5	786	4	US-09-291-417D-23	Sequence 23, Appli
29	42	38.5	787	4	US-09-688-188B-151	Sequence 151, App
30	42	38.5	787	4	US-09-291-417D-151	Sequence 151, App
31	42	38.5	1001	4	US-09-688-188B-31	Sequence 31, Appl
32	42	38.5	1001	4	US-09-291-417D-31	Sequence 31, Appl
33	41	37.6	168	4	US-09-134-001C-4664	Sequence 4664, A
34	41	37.6	453	4	US-09-252-991A-31268	Sequence 31268, A
35	41	37.6	901	3	US-08-936-135-22	Sequence 22, Appl
36	41	37.6	901	3	US-08-936-135-22	Sequence 22, Appl
37	41	37.6	906	3	US-08-936-135-24	Sequence 24, Appl
38	41	37.6	906	3	US-08-936-135-24	Sequence 24, Appl
39	41	37.6	909	3	US-08-936-135-8	Sequence 8, Appli
40	41	37.6	909	3	US-08-936-135-10	Sequence 10, Appl
41	41	37.6	909	3	US-08-936-135-18	Sequence 18, Appl
42	41	37.6	909	4	US-09-439-711C-8	Sequence 8, Appli
43	41	37.6	909	4	US-09-439-711C-10	Sequence 10, Appl
44	41	37.6	909	4	US-09-439-711C-18	Sequence 18, Appl
45	41	37.6	914	3	US-08-936-135-12	Sequence 12, Appl
46	41	37.6	914	4	US-09-439-711C-12	Sequence 12, Appl
47	41	37.6	925	4	US-09-116-473-2	Sequence 2, Appli
48	41	37.6	926	3	US-08-936-135-14	Sequence 14, Appl
49	41	37.6	926	3	US-08-936-135-20	Sequence 20, Appl
50	41	37.6	926	4	US-09-439-711C-14	Sequence 14, Appl
51	41	37.6	926	4	US-09-439-711C-20	Sequence 20, Appl
52	41	37.6	931	3	US-08-936-135-16	Sequence 16, Appl
53	41	37.6	931	4	US-09-439-711C-16	Sequence 16, Appl
54	41	37.6	931	4	US-09-583-638-4	Sequence 4, Appli
55	41	37.6	1805	1	US-07-853-913-2	Sequence 2, Appli
56	40.5	37.2	434	4	US-09-489-039A-8188	Sequence 8188, Ap
57	40	36.7	205	4	US-09-252-991A-18606	Sequence 18606, A
58	40	36.7	275	2	US-08-900-565-4	Sequence 4, Appli
59	40	36.7	275	4	US-09-149-534-4	Sequence 4, Appli
60	40	36.7	295	4	US-09-252-991A-20947	Sequence 20947, A
61	40	36.7	304	4	US-09-763-585-2	Sequence 2, Appli
62	40	36.7	324	4	US-09-763-585-4	Sequence 4, Appli
63	40	36.7	324	4	US-09-252-991A-29070	Sequence 29070, A
64	40	36.7	335	2	US-08-379-556A-10	Sequence 10, Appl
65	40	36.7	407	3	US-08-989-370-5	Sequence 5, Appli
66	40	36.7	407	4	US-09-398-169-5	Sequence 5, Appli
67	40	36.7	951	4	US-09-252-991A-19440	Sequence 19440, A
68	39	35.8	70	4	US-09-621-976-7468	Sequence 7468, Ap
69	39	35.8	146	3	US-08-792-013-7	Sequence 7, Appli
70	39	35.8	147	1	US-08-171-385-5	Sequence 5, Appli
71	39	35.8	147	3	US-08-361-441B-5	Sequence 5, Appli
72	39	35.8	173	4	US-08-252-991A-21967	Sequence 21967, A
73	39	35.8	181	4	US-09-134-001C-2892	Sequence 2892, Ap
74	39	35.8	244	4	US-09-543-681A-7587	Sequence 7587, Ap
75	39	35.8	298	4	US-09-424-349A-7	Sequence 7, Appli
76	39	35.8	418	4	US-09-252-991A-20665	Sequence 20665, A
77	39	35.8	440	4	US-09-077-955-26	Sequence 26, Appl
78	39	35.8	444	4	US-09-424-383-2	Sequence 2, Appli
79	39	35.8	448	4	US-09-328-352-4631	Sequence 4631, Ap
80	39	35.8	448	4	US-09-252-991A-23910	Sequence 23910, A
81	39	35.8	472	4	US-09-252-991A-17633	Sequence 17633, A
82	39	35.8	504	4	US-09-252-991A-29537	Sequence 29537, A
83	39	35.8	524	4	US-09-424-283-1	Sequence 1, Appli
84	39	35.8	660	4	US-09-252-991A-29885	Sequence 29885, A
85	39	35.8	748	4	US-08-688-188B-24	Sequence 24, Appl
86	39	35.8	748	4	US-09-291-417D-24	Sequence 24, Appl
87	39	35.8	1001	3	US-09-060-410-2	Sequence 2, Appli
88	39	35.8	1012	4	US-09-723-458-2	Sequence 2, Appli
89	39	35.8	1012	4	US-08-811-481-16	Sequence 16, Appl
90	39	35.8	1012	4	US-09-876-527-16	Sequence 16, Appl
91	39	35.8	2165	4	US-09-800-729-155	Sequence 155, App
92	38.5	35.3	373	4	US-09-597-771-10	Sequence 10, Appl
93	38.5	35.3	380	4	US-09-254-776B-80	Sequence 80, Appl
94	38.5	35.3	423	4	US-09-328-352-4348	Sequence 4348, Ap
95	38.5	35.3	433	4	US-09-252-991A-27162	Sequence 27162, A
96	38.5	35.3	544	3	US-08-591-685-9	Sequence 9, Appli
97	38.5	35.3	1349	2	US-08-612-734B-2	Sequence 2, Appli
98	38	34.9	16	4	US-09-485-077A-18	Sequence 18, Appl
99	38	34.9	161	4	US-09-489-039A-13372	Sequence 13372, A
100	38	34.9	262	4	US-09-252-991A-25990	Sequence 25990, A

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICHHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.;
; BAROUDY, BAIGIE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/789,262
; FILING DATE: 06-NOV-1991
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2:
; LENGTH: 1091

Query Match 100.0%; Score 109; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 1016 VDDPRSEEDKRFESHIECRK 1035

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A

; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER FILING DATE: 1992-09-18
; EARLIER FILING DATE: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 109; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECKR 20
Db 779 VDDPRSEEDKRFESHIECKR 798

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VDDPRSEEDKRFESHIECRK 20
|||
Db 779 VDDPRSEEDKRFESHIECRK 798

```

RESULT 11
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS,
US-10-135-988-2

```

```
Query Match      100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 VDDPRSEEDKRFESHIECRK 20
779 VDDPRSEEDKRFESHIECRK 798

RESULT 12
US-10-135-988-4
Sequence 4, Application US/10135988
Patent No. 6680060
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PR1
ORGANISM: Attenuated HAV (Pass 35), strain HM
US-10-135-988-4

```
Query Match      100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      1 VDDPRSEEDKRFESHIECRK 20
        |||||
db      779 VDDPRSEEDKRFESHIECRK 798
```

Qy 1 VDDPRSEEDKRFESHIECRK 20
db 779 VDDPRSEEDKRFESHIECRK 798

```

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227

```

TYPE: FRI
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

```
Query Match      100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 VDDPSEEDKRFESHIECRK 20
|||
Db 779 VDDPSEEDKRFESHIECRK 798

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRN
; ORGANISM: Attenuated (4380) HAV, strain HM-17
; US-09-653-499-6

```
Query Match      100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 13
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 2026492US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 109; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||:|||||:|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 14
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 97.2%; Score 106; DB 4; Length 2227;
Best Local Similarity 95.0%; Pred. No. 7.5e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECRK 20
|||||:|||||:|||||
Db 779 VDDPRSEEDKRFESHIECRK 798

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135

; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 74.3%; Score 81; DB 1; Length 839;
Best Local Similarity 82.4%; Pred. No. 0.00019;
Matches 14; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIE 17
|||||:|||||:|||||
Db 779 VDDPRSEEDKRFESHIE 795

RESULT 16
US-09-540-236-3240
; Sequence 3240, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3240
; LENGTH: 112
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3240

Query Match 44.5%; Score 48.5; DB 4; Length 112;
Best Local Similarity 66.7%; Pred. No. 2.4;
Matches 10; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 1 VDDPRSEEDKRFESH 15
| | | | : | : | | | |
Db 50 VDDP-NNSDERFESH 63

RESULT 17

US-09-252-991A-21426
; Sequence 21426, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21426

; LENGTH: 567

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-21426

Query Match

Best Local Similarity 42.2%; Score 46; DB 4; Length 567;

Mismatches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 2 DDPREED--KRFESHIECR 19

| | | | : | : | | | |

Db 367 DDPRAVENRRRLTERLECR 386

RESULT 18

US-08-996-545-2

; Sequence 2, Application US/08996545

; Patent No. 5928898

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.

; APPLICANT: Andrade, Alan C.

; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of

; TITLE OF INVENTION: Aspergillus nidulans

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; STREET: Lilly Corporate Center

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: U.S.

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/996,545

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,872

; REFERENCE/DOCKET NUMBER: X-11766

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334

; TELEFAX: 317-276-2763

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-996-545-2

Query Match 40.8%; Score 44.5; DB 2; Length 1334;
Best Local Similarity 55.0%; Pred. No. 1.4e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 3 DPRSEEDK--RFSHIECR 19

| | | | : | : | | | |

Db 413 DPYSNEGKTLDFEGHIELR 432

RESULT 19

US-09-328-320-2

; Sequence 2, Application US/09328320

; Patent No. 6228615

; GENERAL INFORMATION:

; APPLICANT: Skatrud, Paul L.

; APPLICANT: de Waard, Maarten A.

; APPLICANT: Peery, Robert B.

; APPLICANT: Andrade, Alan C.

; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of

; TITLE OF INVENTION: Aspergillus nidulans

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; STREET: Lilly Corporate Center

; CITY: Indianapolis

; STATE: Indiana

; COUNTRY: U.S.

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/328,320

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/996,545

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Webster, Thomas D.

; REGISTRATION NUMBER: 39,872

; REFERENCE/DOCKET NUMBER: X-11766

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 317-276-3334

; TELEFAX: 317-276-2763

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1334 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-328-320-2

Query Match 40.8%; Score 44.5; DB 3; Length 1334;
Best Local Similarity 55.0%; Pred. No. 1.4e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Qy 3 DPRSEEDK--RFSHIECR 19

| | | | : | : | | | |

Db 413 DPYSNEGKTLDFEGHIELR 432

RESULT 20

US-09-252-991A-30750

; Sequence 30750, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30750
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30750

Query Match 40.4%; Score 44; DB 4; Length 185;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 4 PRSEEDKRFESHIECR 19
| | | : | : | : |
Db 92 PASEAEQRYEVQIVCR 107

RESULT 21
US-09-252-991A-31407
; Sequence 31407, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31407
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31407

Query Match 40.4%; Score 44; DB 4; Length 496;
Best Local Similarity 38.9%; Pred. No. 59;
Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECR 19
| | | : | : | : |
Db 362 EDPAESHRGQGRTECR 379

RESULT 22
US-09-328-352-5208
; Sequence 5208, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5208

; LENGTH: 818
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5208

Query Match 39.9%; Score 43.5; DB 4; Length 818;
Best Local Similarity 47.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 3 DPRSEEDKRFESHIECR 19
| | | : | : | : |
Db 266 DPKLTKDQSPQ-HLTCT 281

RESULT 23
US-08-887-798-2
; Sequence 2, Application US/08887798
; Patent No. 5922556
; GENERAL INFORMATION:
; APPLICANT: Mayeux, Richard
; APPLICANT: Graziano, Joseph H.
; APPLICANT: Freyer, Greg
; TITLE OF INVENTION: PARKINSON'S DISEASE TESTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/887,798
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/51949/JPW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 780 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-887-798-2

Query Match 39.4%; Score 43; DB 2; Length 780;
Best Local Similarity 42.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHIECR 19
| | | : | : | : |
Db 74 LDDPASQEIERRKSYLR 92

RESULT 24
5273901-7
; Patent No. 5273901
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: GENETICALLY ENGINEERED COCCIDIOSIS
; SPOROZOITE 21.5 KB ANTIGEN, AC-6B
; NUMBER OF SEQUENCES: 11

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/581.693
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 215.162
; FILING DATE: 05-JUL-1988
; APPLICATION NUMBER: 746.520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627.811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:7:
; LENGTH:180
5273901-7

Query Match      38.5%; Score 42; DB 6; Length 180;
Best Local Similarity 56.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIE 17
   ||| ||||| |
DB 3 EDKREEDKREEEEE 18

RESULT 25
5482709-6
; Patent No. 5482709
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH
; ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148.432
; FILING DATE: 08-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 581.693
; FILING DATE: 12-SEP-1990
; APPLICATION NUMBER: 215.162
; FILING DATE: 05-JUL-1989
; APPLICATION NUMBER: 746.520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627.811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:6:
; LENGTH: 180
5482709-6

Query Match      38.5%; Score 42; DB 6; Length 180;
Best Local Similarity 56.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIE 17
   ||| ||||| |
DB 3 EDKREEDKREEEEE 18

RESULT 26
US-08-559-492-5
; Sequence 5, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/581.693
; FILING DATE: 12-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 215.162
; FILING DATE: 05-JUL-1988
; APPLICATION NUMBER: 746.520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627.811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:7:
; LENGTH:180
5273901-7

Query Match      38.5%; Score 42; DB 6; Length 180;
Best Local Similarity 56.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIE 17
   ||| ||||| |
DB 3 EDKREEDKREEEEE 18

RESULT 25
5482709-6
; Patent No. 5482709
; APPLICANT: JACOBSON, JAMES W.; STRAUSBERG, ROBERT L.; WILSON,
; SUSAN D.; POPE, SHARON H.; STRAUSBERG, SUSAN L.; RUFF, MICHAEL D.;
; AUGUSTINE, PATRICIA C.; DANFORTH, HARRY D.
; TITLE OF INVENTION: EIMERIA ANTIGENIC COMPOSITION WHICH
; ELICITS ANTIBODIES AGAINST AVIAN COCCIDIOSIS
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148.432
; FILING DATE: 08-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 581.693
; FILING DATE: 12-SEP-1990
; APPLICATION NUMBER: 215.162
; FILING DATE: 05-JUL-1989
; APPLICATION NUMBER: 746.520
; FILING DATE: 19-JUN-1985
; APPLICATION NUMBER: 627.811
; FILING DATE: 05-JUL-1984
; SEQ ID NO:6:
; LENGTH: 180
5482709-6

Query Match      38.5%; Score 42; DB 6; Length 180;
Best Local Similarity 56.2%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIE 17
   ||| ||||| |
DB 3 EDKREEDKREEEEE 18

RESULT 26
US-08-559-492-5
; Sequence 5, Application US/08559492
; Patent No. 5843884
; GENERAL INFORMATION:
; APPLICANT: Sims, Peter J.
; TITLE OF INVENTION: C9 Complement Inhibitor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center, 1201 West Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559.492
; FILING DATE: 15-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRF154
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-873-8794
; TELEFAX: 404-873-8795
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-559-492-5

Query Match      38.5%; Score 42; DB 2; Length 560;
Best Local Similarity 47.1%; Pred. No. 1.4e+02;
Matches 8; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 DPRSEEDKRFESHIECR 19
   ||| | |||||
DB 29 DPELTSSEGSASHIDCR 45

RESULT 27
US-09-688-188B-23
; Sequence 23, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 23
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-23

Query Match      38.5%; Score 42; DB 4; Length 786;
Best Local Similarity 40.0%; Pred. No. 2e+02;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 6 SEEDKRFESHIECRX 20
   |||:|:|:|:|:|
DB 315 SNEKKFQOHTQAQQ 329

RESULT 28
US-09-291-417D-23
; Sequence 23, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
```

```

; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHITE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 786
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-417D-23

```

Query Match 38.5%; Score 42; DB 4; Length 786;
Best Local Similarity 40.0%; Pred.No.2e+02;
Matches 6; Conservative 6; Mismatches 3; Indels

Qy 6 SEEDKRFESHIECRK 20
315 SNEEKKFOOHIOAOO 329

RESULT 29

```

US-09-688-188B-151
; Sequence 151, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0328
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 151
; LENGTH: 787
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-688-188B-151

```

Query Match 38.5%; Score 42; DB 4; Length 787;
Best Local Similarity 40.0%; Pred. No. 2e+02;
Matches 6; Conservative 6; Mismatches 3; Indels

Qy 6 SEEDKRFESHIECRK 20
| | | | | : | : | : | :
Db 316 SNEEKFOOHIOAOO 330

RESULT 30

US-09-291-417D-151
 ; Sequence 151, Application US/09291417D
 ; Patent No. 6680170
 ; GENERAL INFORMATION:
 ; APPLICANT: PLOWMAN, GREGORY
 ; APPLICANT: MARTINEZ, RICARDO
 ; APPLICANT: WHYTE, DAVID
 ; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
 ; FILE REFERENCE: 038602/0329
 ; CURRENT APPLICATION NUMBER: US/09/291.417D
 ; CURRENT FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 60/081,784
 ; PRIOR FILING DATE: 1998-04-14
 ; NUMBER OF SEQ ID NOS: 155

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:37:37 ; Search time 34.25 Seconds
(without alignments)

162.083 Million cell updates/sec

Title: US-09-171-432A-38
Perfect score: 109
Sequence: 1 VDDPRSEDKRFESHIECKR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US02_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/PCTUS_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	109	100.0	20	10	US-09-171-432A-38
2	109	100.0	2227	13	US-10-135-988-2
3	109	100.0	2227	13	US-10-135-988-4
4	109	100.0	2227	13	US-10-135-988-6
5	106	97.2	836	14	US-10-272-459-40
6	106	97.2	980	14	US-10-272-459-41
7	106	97.2	2227	9	US-09-929-955-12
8	106	97.2	2227	13	US-10-104-966-12
9	69	63.3	20	10	US-09-171-432A-37
10	66	60.6	302	14	US-10-272-459-44
11	51	45.8	1548	15	US-10-369-493-6347
12	49	45.0	1661	10	US-09-842-758-42
13	49	45.0	1661	12	US-10-174-333-42
14	47	43.1	352	12	US-10-425-114-65499
15	46	42.2	1663	10	US-09-842-758-43
16	46	42.2	1663	10	US-09-842-758-43
17	46	42.2	1663	10	US-09-842-758-43
18	46	42.2	1663	10	US-09-842-758-43
19	46	42.2	1663	10	US-09-842-758-43
20	46	42.2	1663	10	US-09-842-758-43
21	46	42.2	1663	10	US-09-842-758-43
22	46	42.2	1663	10	US-09-842-758-43
23	46	42.2	1663	10	US-09-842-758-43
24	46	42.2	1663	10	US-09-842-758-43
25	46	42.2	1663	10	US-09-842-758-43
26	46	42.2	1663	10	US-09-842-758-43
27	46	42.2	1663	10	US-09-842-758-43
28	46	42.2	1663	10	US-09-842-758-43
29	46	42.2	1663	10	US-09-842-758-43
30	46	42.2	1663	10	US-09-842-758-43
31	46	42.2	1663	10	US-09-842-758-43
32	46	42.2	1663	10	US-09-842-758-43
33	46	42.2	1663	10	US-09-842-758-43
34	46	42.2	1663	10	US-09-842-758-43
35	46	42.2	1663	10	US-09-842-758-43
36	46	42.2	1663	10	US-09-842-758-43
37	46	42.2	1663	10	US-09-842-758-43
38	46	42.2	1663	10	US-09-842-758-43
39	46	42.2	1663	10	US-09-842-758-43
40	46	42.2	1663	10	US-09-842-758-43
41	46	42.2	1663	10	US-09-842-758-43
42	46	42.2	1663	10	US-09-842-758-43
43	46	42.2	1663	10	US-09-842-758-43
44	46	42.2	1663	10	US-09-842-758-43
45	46	42.2	1663	10	US-09-842-758-43
46	46	42.2	1663	10	US-09-842-758-43
47	46	42.2	1663	10	US-09-842-758-43
48	46	42.2	1663	10	US-09-842-758-43
49	46	42.2	1663	10	US-09-842-758-43
50	46	42.2	1663	10	US-09-842-758-43
51	46	42.2	1663	10	US-09-842-758-43
52	46	42.2	1663	10	US-09-842-758-43
53	46	42.2	1663	10	US-09-842-758-43
54	46	42.2	1663	10	US-09-842-758-43
55	46	42.2	1663	10	US-09-842-758-43
56	46	42.2	1663	10	US-09-842-758-43
57	46	42.2	1663	10	US-09-842-758-43
58	46	42.2	1663	10	US-09-842-758-43
59	46	42.2	1663	10	US-09-842-758-43
60	46	42.2	1663	10	US-09-842-758-43
61	46	42.2	1663	10	US-09-842-758-43
62	46	42.2	1663	10	US-09-842-758-43
63	46	42.2	1663	10	US-09-842-758-43
64	46	42.2	1663	10	US-09-842-758-43
65	46	42.2	1663	10	US-09-842-758-43
66	46	42.2	1663	10	US-09-842-758-43
67	46	42.2	1663	10	US-09-842-758-43
68	46	42.2	1663	10	US-09-842-758-43
69	46	42.2	1663	10	US-09-842-758-43
70	46	42.2	1663	10	US-09-842-758-43
71	46	42.2	1663	10	US-09-842-758-43
72	46	42.2	1663	10	US-09-842-758-43
73	46	42.2	1663	10	US-09-842-758-43
74	46	42.2	1663	10	US-09-842-758-43
75	46	42.2	1663	10	US-09-842-758-43
76	46	42.2	1663	10	US-09-842-758-43
77	46	42.2	1663	10	US-09-842-758-43
78	46	42.2	1663	10	US-09-842-758-43
79	46	42.2	1663	10	US-09-842-758-43
80	46	42.2	1663	10	US-09-842-758-43
81	46	42.2	1663	10	US-09-842-758-43
82	46	42.2	1663	10	US-09-842-758-43
83	46	42.2	1663	10	US-09-842-758-43
84	46	42.2	1663	10	US-09-842-758-43
85	46	42.2	1663	10	US-09-842-758-43
86	46	42.2	1663	10	US-09-842-758-43
87	46	42.2	1663	10	US-09-842-758-43
88	46	42.2	1663	10	US-09-842-758-43
89	46	42.2	1663	10	US-09-842-758-43
90	46	42.2	1663	10	US-09-842-758-43
91	46	42.2	1663	10	US-09-842-758-43
92	46	42.2	1663	10	US-09-842-758-43
93	46	42.2	1663	10	US-09-842-758-43
94	46	42.2	1663	10	US-09-842-758-43
95	46	42.2	1663	10	US-09-842-758-43
96	46	42.2	1663	10	US-09-842-758-43
97	46	42.2	1663	10	US-09-842-758-43
98	46	42.2	1663	10	US-09-842-758-43
99	46	42.2	1663	10	US-09-842-758-43
100	46	42.2	1663	10	US-09-842-758-43

Sequence 43, Appl
Sequence 189646,
Sequence 272156,
Sequence 6516, Ap
Sequence 959, App
Sequence 2, Appl
Sequence 175273,
Sequence 175273,
Sequence 6, Appl
Sequence 160177,
Sequence 18071, A
Sequence 3307, Ap
Sequence 5538, A
Sequence 1498, Ap
Sequence 8, Appl
Sequence 197989,
Sequence 5558, Ap
Sequence 8, Appl
Sequence 36, Appl
Sequence 173280,
Sequence 1936, Ap
Sequence 203313,
Sequence 153364,
Sequence 1430, Ap
Sequence 1869, Ap
Sequence 77695, A
Sequence 193, App
Sequence 23, Appl
Sequence 31, Appl
Sequence 254, App
Sequence 18, Appl
Sequence 2949, Ap
Sequence 122, App
Sequence 36, Appl
Sequence 2729, Ap
Sequence 5658, A
Sequence 78017, A
Sequence 7161, Ap
Sequence 22799, A
Sequence 5915, Ap
Sequence 5916, Ap
Sequence 5916, Ap
Sequence 4, Appl
Sequence 50816, A
Sequence 4, Appl
Sequence 8, Appl
Sequence 17, App
Sequence 4, Appl
Sequence 4, Appl
Sequence 3622, Ap
Sequence 87, Appl
Sequence 5013, Ap
Sequence 271955,
Sequence 45342, A
Sequence 7, Appl
Sequence 452, App
Sequence 2317, Ap
Sequence 39, Appl
Sequence 212731,
Sequence 1368, Ap
Sequence 48171, A
Sequence 28603, A
Sequence 342, App
Sequence 160385,
Sequence 4, Appl
Sequence 54392, A
Sequence 158843,
Sequence 153076,
Sequence 45209, A
Sequence 45, Appl
Sequence 7160, Ap
Sequence 7000, Ap
Sequence 40451, A
Sequence 8492, Ap

Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDI, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
Query Match 100.0%; Score 109; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDKRFESHIECKR 798
RESULT 5
US-10-272-459-40
Sequence 40, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHUANES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
TITLE OF INVENTION: PROTEINS AND USES THEREOF
FILE REFERENCE: PP17955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 40
LENGTH: 836
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-40
Query Match 97.2%; Score 106; DB 14; Length 836;
Best Local Similarity 95.0%; Pred. No. 3e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDKRFESHIECKR 798
RESULT 6
US-10-272-459-41
Sequence 41, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHUANES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
TITLE OF INVENTION: PROTEINS AND USES THEREOF
FILE REFERENCE: PP17955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15

NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 41
LENGTH: 980
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
Query Match 97.2%; Score 106; DB 14; Length 980;
Best Local Similarity 95.0%; Pred. No. 3.6e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDKRFESHIECKR 798
RESULT 7
US-09-929-955-12
Sequence 12, Application US/09929955
Patent No. US20020136740A1
GENERAL INFORMATION:
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 2227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
Query Match 97.2%; Score 106; DB 9; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VDDPRSEEDKRFESHIECKR 20
DB 779 VDDPRSEEDKRFESHIECKR 798
RESULT 8
US-10-104-966-12
Sequence 12, Application US/10104966
Publication No. US20020155124A1
GENERAL INFORMATION:
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: TRIPEP.23AUSC1
CURRENT APPLICATION NUMBER: US/10/104,966
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29

```
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          97.28; Score 106; DB 13; Length 2227;
Best Local Similarity 95.08; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 VDDPRSEEDKRFESHIECRK 20
      |||||.....|
Db      779 VDDPRSEEDKRFESHIECRK 798
      |||||.....|

RESULT 9
US-09-171-432A-37
; Sequence 37, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESS: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1-20
; OTHER INFORMATION: /label=YK-1313
US-09-171-432A-37

Query Match          63.33; Score 69; DB 10; Length 20;
Best Local Similarity 100.08; Pred. No. 0.0019;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VDDPRSEEDKRFPE 13
      |||||.....|
Db      1 VDDPRSEEDKRFPE 13
      |||||.....|

US-10-272-459-44
; Sequence 44, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHAUNTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272.459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: Of 33.3 kDa
US-10-272-459-44

Query Match          60.66; Score 66; DB 14; Length 302;
Best Local Similarity 92.33; Pred. No. 0.095;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 VDDPRSEEDKRFPE 13
      |||||.....|
Db      290 VDDPRSEEDKRFPE 302
      |||||.....|

RESULT 11
US-10-369-493-6347
; Sequence 6347, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiaofeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6347
; LENGTH: 1548
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6347

Query Match          46.88; Score 51; DB 15; Length 1548;
Best Local Similarity 60.08; Pred. No. 90;
Matches 9; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      2 DDPRSEEDKRFESHI 16
      ::|||:::|
Db      479 ENPRSEEDRNYESTI 493
      ::|||:::|

RESULT 12
US-09-842-758-42
; Sequence 42, Application US/09842758
; Publication No. US20030083244A1
; GENERAL INFORMATION:
```

```

; Publication No. US20040029220A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shinkets, Richard A
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E
; APPLICANT: Gangolli, Esha A
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Alsobrook, John P
; APPLICANT: Anderson, David W
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Li, Li
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-783 CIP1
; CURRENT APPLICATION NUMBER: US/10/174,333
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/193,664
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/194,614
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,063
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,066
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,067
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,068
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,070
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,510
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/219,855
; PRIOR FILING DATE: 2000-07-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 42
; LENGTH: 1661
; TYPE: PRT
; ORGANISM: Sus scrofa
; US-10-174-333-42

Query Match 45.0%; Score 49; DB 10; Length 1661;
Best Local Similarity 42.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIECRK 20
Db 1572 DEVQVGERRFISHIKRE 1590
| : : : : : |
| : : : : : |

RESULT 13
US-10-174-333-42
; Sequence 42, Application US/10174333
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Fernandes, Elma R.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shinkets, Richard A
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Majumder, Kumud
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Burgess, Catherine E
; APPLICANT: Gangolli, Esha A
; APPLICANT: Smithson, Glenda
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier, Raymond J
; APPLICANT: Grosse, William M
; APPLICANT: Edward, Szekeres S
; APPLICANT: Alsobrook II, John P
; APPLICANT: Alsobrook II, John P
; TITLE OF INVENTION: NO. US20030083244A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-783
; CURRENT APPLICATION NUMBER: US/09/842,758
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/200,158
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 60/200,613
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,780
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/201,006
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,007
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,236
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,238
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/201,186
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: 60/201,474
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/201,508
; PRIOR FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 60/220,591
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/232,678
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/263,217
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 60/265,160
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 1661
; TYPE: PRT
; ORGANISM: Sus scrofa
; US-09-842-758-42

Query Match 45.0%; Score 49; DB 10; Length 1661;
Best Local Similarity 42.1%; Pred. No. 1.9e+02;
Matches 8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 DDPSEEDKRFESHIECRK 20
Db 1572 DEVQVGERRFISHIKRE 1590
| : : : : : |
| : : : : : |

RESULT 13
US-10-174-333-42
; Sequence 42, Application US/10174333
; Publication No. US20040034888A1
; GENERAL INFORMATION:
```

APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 65499
LENGTH: 352
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4766-014-A10_F11.pep
US-10-425-114-65499

Query Match 43.1%; Score 47; DB 12; Length 352;
Best Local Similarity 40.9%; Pred. No. 75;
Matches 9; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 1 VDDPRSEEDKRFES--HIECRK 20
DB 235 IDTGRRDRDKAFALLHLQCPK 256

RESULT 15
US-09-842-758-43
Sequence 43, Application US/09842758
Publication No. US20030083244A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Fernandes, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Shimkets, Richard A
APPLICANT: Malyankar, Uriel M
APPLICANT: Boldog, Ferenc L
APPLICANT: Zethusen, Bryan D
APPLICANT: Spytek, Kimberly A
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine E
APPLICANT: Gangolli, Esha A
APPLICANT: Smithson, Glennda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R
APPLICANT: Taupier, Raymond J
APPLICANT: Grosse, William M
APPLICANT: Edward, Szekeres S
APPLICANT: Alsobrook II, John P
TITLE OF INVENTION: No. US20030083244A1el Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 15966-783
CURRENT APPLICATION NUMBER: US/09/842,758
CURRENT FILING DATE: 2001-04-25
PRIOR APPLICATION NUMBER: 60/200,158
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 60/200,613
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,780
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/201,006
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,007
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,236
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/201,238
PRIOR FILING DATE: 2000-05-01

PRIOR APPLICATION NUMBER: 60/201,186
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: 60/201,474
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/201,508
PRIOR FILING DATE: 2000-05-03
PRIOR APPLICATION NUMBER: 60/220,591
PRIOR FILING DATE: 2000-07-25
PRIOR APPLICATION NUMBER: 60/232,678
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 60/263,217
PRIOR FILING DATE: 2001-01-22
PRIOR APPLICATION NUMBER: 60/265,160
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 113
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 43
LENGTH: 1663
TYPE: PRT
ORGANISM: Mus musculus
US-09-842-758-43

Query Match 42.2%; Score 46; DB 10; Length 1663;
Best Local Similarity 38.9%; Pred. No. 5.4e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECR 19
DB 1574 DEVQAGQQRKFISHIKR 1591

RESULT 16
US-10-174-333-43
Sequence 43, Application US/10174333
Publication No. US20040029220A1
GENERAL INFORMATION:
APPLICANT: Vernet, Corine A. M.
APPLICANT: Fernandes, Elma R.
APPLICANT: Gerlach, Valerie
APPLICANT: Malyankar, Uriel M.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Zethusen, Bryan D.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Majumder, Kumud
APPLICANT: Tchernev, Velizar T.
APPLICANT: Padigaru, Muralidhara
APPLICANT: Patturajan, Meera
APPLICANT: Burgess, Catherine E.
APPLICANT: Gangolli, Esha A.
APPLICANT: Smithson, Glennda
APPLICANT: Rastelli, Luca
APPLICANT: MacDougall, John R.
APPLICANT: Taupier, Raymond J.
APPLICANT: Grosse, William M.
APPLICANT: Szekeres, Edward S.
APPLICANT: Alsobrook, John P.
APPLICANT: Anderson, David W.
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Li, Li
APPLICANT: Zhong, Mei
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 15966-783 CIP1
CURRENT APPLICATION NUMBER: US/10/174,333
CURRENT FILING DATE: 2002-06-18
PRIOR APPLICATION NUMBER: 60/193,664
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 60/194,614
PRIOR FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: 60/195,063
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,066
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 60/195,067

```

; PRIOR FILLING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILLING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,069
; PRIOR FILLING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,070
; PRIOR FILLING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,510
; PRIOR FILLING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/219,855
; PRIOR FILLING DATE: 2000-07-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 43
; LENGTH: 1663
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-174-333-43

```

```

Query Match      42.2%; Score 46; DB 12; Length 1653;
Best Local Similarity 38.3%; Pred. NO. 5.4e+02;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      2 DDPRSSEDKRPFESHIECR 19
      |: : : : |||
      |: : : : |||

Db      1574 DEVQAGQQRKF-SHIKCR 1591

```

```

RESULT 17
US-10-424-599-189646
; Sequence 189646, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovacic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 189646
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142256C.1.pep
US-10-424-599-189646

```

```

Query Match      41.3%; Score 45; DB 12; Length 57;
Best Local Similarity 69.2%; Pred. NO. 22;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      6 SEEDKRFESHIEC 18
      | | | | |
Db      10 SRVQKRFSLSHIEC 22

```

RESULT 18
US-10-424-599-272156
; Sequence 272156, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B

```
; TYPE: PRT
; ORGANISM: Archaeoglobus fulgidus
US-10-369-493-959

Query Match      41.3%; Score 45; DB 15; Length 586;
Best Local Similarity 53.3%; Pred. No. 2.5e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      2 DDPRSEEDKRFESHI 16
DB      356 DEPDEEKDKLVH 370
      :|:|:|:|:|:|
      :|:|:|:|:|:|

RESULT 21
US-09-758-828-2
; Sequence 2, Application US/09758828
; Patent No. US20020164772A1
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atRD of
; Aspergillus nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/758,828
; FILING DATE: 11-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-758-828-2

Query Match      40.8%; Score 44.5; DB 9; Length 1334;
Best Local Similarity 55.0%; Pred. No. 7.2e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY      3 DPRSEEDK---RFESHICR 19
DB      413 DPYNSGKTLDFHFEHIELR 432
      |||:|:|:|:|:|:|
      |||:|:|:|:|:|:|

RESULT 22
US-10-424-599-175273
; Sequence 175273, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175273
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129290C.1.pep
US-10-424-599-175273

Query Match      40.4%; Score 44; DB 12; Length 91;
Best Local Similarity 36.0%; Pred. No. 51;
Matches 9; Conservative 5; Mismatches 5; Indels 6; Gaps 2;

QY      1 VDDPRSEEDKRFESHI-----ECRK 20
DB      16 MEDPK-ERDRRFREHYRVKRCCK 39
      :|:|:|:|:|:|:|
      :|:|:|:|:|:|:|

RESULT 23
US-10-424-599-175274
; Sequence 175274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 175274
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_129291C.1.pep
US-10-424-599-175274

Query Match      40.4%; Score 44; DB 12; Length 91;
Best Local Similarity 36.0%; Pred. No. 51;
Matches 9; Conservative 5; Mismatches 5; Indels 6; Gaps 2;

QY      1 VDDPRSEEDKRFESHI-----ECRK 20
DB      16 MEDPK-ERDRRFREHYRVKRCCK 39
      :|:|:|:|:|:|:|
      :|:|:|:|:|:|:|

RESULT 24
US-10-427-224-6
; Sequence 6, Application US/10427224
; Publication No. US20040033607A1
; GENERAL INFORMATION:
; APPLICANT: Van No. US20040033607A1ker, Steven R.
; APPLICANT: Zhang, Hua
; TITLE OF INVENTION: Plant Vernalization Independence (VIP) Genes, Proteins, and
; FILE REFERENCE: MSU-08107
; CURRENT APPLICATION NUMBER: US/10/427,224
; CURRENT FILING DATE: 2003-05-01
; PRIOR APPLICATION NUMBER: 60/376,765
```

```

; NAME/KEY: unsure
; LOCATION: (1)..(223)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-18071

Query Match 39.4%; Score 43; DB 15; Length 223;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECRK 20
Db 11 EEPFAGDKRAEATHTHR 29

RESULT 27
US-10-369-493-3307
; Sequence 3307, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3307
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3307

Query Match 39.4%; Score 43; DB 15; Length 422;
Best Local Similarity 47.6%; Pred. No. 3.6e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 2 DDPRSEEDKRFESHIEC 18
Db 217 DEERAERGFPFEGSFESQIEC 237

RESULT 28
US-10-425-114-55238
; Sequence 55238, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55238
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: UC-ZMFLB73019H10_FLI.pcp
US-10-425-114-55238

Query Match 39.4%; Score 43; DB 12; Length 535;
; FEATURE:
; ORGANISM: SPHINGOMONAS
; NAME/KEY: unsure
; LOCATION: (1)..(223)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-18071

Query Match 39.9%; Score 43.5; DB 12; Length 625;
Best Local Similarity 45.5%; Pred. No. 4.6e+02;
Matches 10; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

QY 1 VDDPRSEEDKRFES---HIECR 19
Db 214 LDDIIFEDPQYSEAEHVEAR 235

RESULT 25
US-10-424-599-166177
; Sequence 166177, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166177
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121073C.1.pcp
US-10-424-599-166177

Query Match 39.4%; Score 43; DB 12; Length 96;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHI 16
Db 19 VDNPSAVDDLAFHSHI 34

RESULT 26
US-10-369-493-18071
; Sequence 18071, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18071
; LENGTH: 223
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
; NAME/KEY: unsure
; LOCATION: (1)..(223)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-18071

Query Match 39.4%; Score 43; DB 15; Length 223;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECRK 20
Db 11 EEPFAGDKRAEATHTHR 29

RESULT 27
US-10-369-493-3307
; Sequence 3307, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3307
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3307

Query Match 39.4%; Score 43; DB 15; Length 422;
Best Local Similarity 47.6%; Pred. No. 3.6e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 2 DDPRSEEDKRFESHIEC 18
Db 217 DEERAERGFPFEGSFESQIEC 237

RESULT 28
US-10-425-114-55238
; Sequence 55238, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 55238
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Zea mays
; OTHER INFORMATION: Clone ID: UC-ZMFLB73019H10_FLI.pcp
US-10-425-114-55238

Query Match 39.4%; Score 43; DB 12; Length 535;
; FEATURE:
; ORGANISM: SPHINGOMONAS
; NAME/KEY: unsure
; LOCATION: (1)..(223)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-18071

Query Match 39.9%; Score 43.5; DB 12; Length 625;
Best Local Similarity 45.5%; Pred. No. 4.6e+02;
Matches 10; Conservative 4; Mismatches 5; Indels 3; Gaps 1;

QY 1 VDDPRSEEDKRFES---HIECR 19
Db 214 LDDIIFEDPQYSEAEHVEAR 235

RESULT 25
US-10-424-599-166177
; Sequence 166177, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166177
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_121073C.1.pcp
US-10-424-599-166177

Query Match 39.4%; Score 43; DB 12; Length 96;
Best Local Similarity 50.0%; Pred. No. 75;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 VDDPRSEEDKRFESHI 16
Db 19 VDNPSAVDDLAFHSHI 34

RESULT 26
US-10-369-493-18071
; Sequence 18071, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 18071
; LENGTH: 223
; TYPE: PRT
; ORGANISM: SPHINGOMONAS
; NAME/KEY: unsure
; LOCATION: (1)..(223)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-18071

Query Match 39.4%; Score 43; DB 15; Length 223;
Best Local Similarity 36.8%; Pred. No. 1.8e+02;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 DDPRSEEDKRFESHIECRK 20
Db 11 EEPFAGDKRAEATHTHR 29

RESULT 27
US-10-369-493-3307
; Sequence 3307, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3307
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Neurospora crassa
US-10-369-493-3307

Query Match 39.4%; Score 43; DB 15; Length 422;
Best Local Similarity 47.6%; Pred. No. 3.6e+02;
Matches 10; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 2 DDPR
```


Best Local Similarity 70.0%; Pred. No. 4.6e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
Qy 2 DDPSEEDKR 11
:||||:|
Db 496 DNPSEQDKR 505

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
Qy 1 VDDPSEEDKRFESHIECR 19
:||||:|:|:|:|:|:|
Db 74 LDDPASQETIERGKSYLR 92

Search completed: May 11, 2004, 13:44:48
Job time : 35.25 secs

RESULT 29
US-10-389-566-1498
; Sequence 1498, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1498
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Fusobacterium nucleatum
US-10-389-566-1498

Query Match 39.4%; Score 43; DB 16; Length 607;
Best Local Similarity 41.2%; Pred. No. 5.2e+02;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
Qy 3 DDPSEEDKRFESHIECR 19
:||||:|:|
Db 491 EAAEEDKKFQELVEAR 507

RESULT 30
US-10-168-066-8
; Sequence 8, Application US/10168066
; Publication No. US20030087268A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: LJ, Dying Aina M.
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: HUMAN LYASES AND ASSOCIATED PROTEINS
; FILE REFERENCE: PF 0759 PCT
; CURRENT APPLICATION NUMBER: US/10/168,066
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/172,307
; PRIOR FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 8
; LENGTH: 780
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030087268A1 3130234
US-10-168-066-8

Query Match 39.4%; Score 43; DB 14; Length 780;
Best Local Similarity 42.1%; Pred. No. 6.8e+02;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:50:08 ; Search time 10 seconds
(without alignments)
192.383 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107

Sequence: 1 SHIECRKPYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	56	PQ0431	genome polyprotein
2	107	100.0	341	S04137	genome polyprotein
3	107	100.0	852	GNHYHA	genome polyprotein
4	107	100.0	1358	A03905	genome polyprotein
5	107	100.0	2227	GNHYHM	genome polyprotein
6	107	100.0	2227	GNHYHR	genome polyprotein
7	107	100.0	2227	GNHYMK	genome polyprotein
8	107	100.0	2227	GNHYHB	genome polyprotein
9	105	98.1	56	PQ0427	genome polyprotein
10	104	97.2	56	PQ0429	genome polyprotein
11	102	95.3	319	JH0135	genome polyprotein
12	100	93.5	56	PQ0430	genome polyprotein
13	97	90.7	56	PQ0432	genome polyprotein
14	96	89.7	56	PQ0428	genome polyprotein
15	95	88.8	56	PQ0434	genome polyprotein
16	87	81.3	55	PQ0433	genome polyprotein
17	86	80.4	56	PQ0436	genome polyprotein
18	86	80.4	2230	GNVNSA	genome polyprotein
19	84.5	79.0	55	PQ0435	genome polyprotein
20	84.5	79.0	839	GNVNS2	genome polyprotein
21	81	75.7	56	PQ0437	genome polyprotein
22	46	43.0	420	S62541	hypothetical prote
23	44	41.1	374	C84C40	hypothetical prote
24	43.5	40.7	1663	C35T	complement C3 prec
25	43	40.2	332	P95568	probable lipase, 2
26	43	40.2	339	T18926	hypothetical prote
27	43	40.2	343	AH1823	30S ribosomal prot
28	43	40.2	346	T27896	hypothetical prote
29	43	40.2	391	VHNZ3	nucleocapsid prote

30 43 40.2 554 2 T49917
31 42.5 39.7 140 2 E64364
32 42 39.3 121 2 T17708
33 42 39.3 214 2 E82679
34 42 39.3 286 2 S07532
35 42 39.3 329 2 T25067
36 42 39.3 494 2 S62902
37 42 39.3 723 2 D71091
38 42 39.3 861 2 B84963
39 42 39.3 1087 2 T16876
40 42 39.3 1157 2 T43259
41 42 39.3 1270 2 T22615
42 41.5 38.8 286 2 H87506
43 41.5 38.8 606 2 A12649
44 41.5 38.8 606 2 H97431
45 41 38.3 144 2 H96511

hypothetical prote
hypothetical prote
hypothetical prote
chaperone XF1452 I
puff 11/9-1 protei
hypothetical prote
legumin 2 precurs
methionine-tRNA li
DNA topoisomerase
hypothetical prote
pyruvate (flavodox
hypothetical prote
phosphogluconate d
phosphogluconate d
AP2 domain contain

ALIGNMENTS

RESULT 1

PQ0431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wide
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geog
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0431
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 107; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
|||
DB 29 SHIECRKPYKELRLEVGKQR 48
|||

RESULT 2

S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A:Reference number: S04137; MUID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match 100.0%; Score 107; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
|||

Db 296 SHIECRPKYKELRLEVGVKQR 315

RESULT 3

GNVYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C/Accession: A03904
R/Lineage: D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.
J. Virol. 54, 247-255, 1985
A/Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A/Reference number: A03904; MUID:85185648; PMID:2985793
A/Accession: A03904
A/Molecule type: Genomic RNA
A/Residues: 1-852 <LIN>
A/Cross-references: EMBL:M10033; NID:G329592; PID:AAA45470.1; PID:G329593
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein
F/1-245/Product: coat protein 1A #status predicted <CIA>
F/246-491/Product: coat protein 1B #status predicted <CIB>
F/492-836/Product: coat protein 1C #status predicted <CIC>
F/837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 107; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 6.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGVKQR 20

Db 792 SHIECRPKYKELRLEVGVKQR 811

RESULT 4

A03905
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein 3
C/Species: human hepatitis A virus
C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C/Accession: A03905
R/Baroudy, B.M.; Ticehurst, J.R.; Mile, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstone
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A/Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A/Reference number: A03905; MUID:85166289; PMID:2984684
A/Accession: A03905
A/Molecule type: Genomic RNA
A/Residues: 1-1358 <BAR>
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; polyprotein
F/1-245/Product: coat protein 1A #status predicted <CIA>
F/246-491/Product: coat protein 1B #status predicted <CIB>
F/492-836/Product: coat protein 1C #status predicted <CIC>
F/837-854/Product: core protein 2A (fragment) #status predicted <C2A>
F/855-1358/Product: protein 3D (fragment) #status predicted <C3D>

Query Match 100.0%; Score 107; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGVKQR 20

Db 792 SHIECRPKYKELRLEVGVKQR 811

RESULT 5

GNVYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 3
B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)

C/Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C/Accession: A25981
R/Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A/Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with c
A/Reference number: A25981; MUID:87061253; PMID:3023706
A/Accession: A25981
A/Molecule type: Genomic RNA
A/Residues: 1-2227 <COH>

A/Cross-references: EMBL:M44707; NID:G329582; PID:AAA45465.1; PID:G329583
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
F/1-23/Product: coat protein 1A #status predicted <VP4>
F/24-245/Product: coat protein 1B #status predicted <VP2>
F/246-491/Product: coat protein 1C #status predicted <VP3>
F/492-791/Product: coat protein 1D #status predicted <VP1>
F/792-980/Product: core protein 2A #status predicted <C2A>
F/981-1087/Product: core protein 2B #status predicted <C2B>
F/1088-1422/Product: core protein 2C #status predicted <C2C>
F/1423-1496/Product: protein 3A #status predicted <C3A>
F/1497-1519/Product: protein 3B #status predicted <C3B>
F/1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F/1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGVKQR 20

Db 792 SHIECRPKYKELRLEVGVKQR 811

RESULT 6

GNVYHR
genome polyprotein - human hepatitis A virus
N/Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pr
NA polymerase (EC 2.7.7.48), protein 3D
C/Species: human hepatitis A virus
A/Note: host Homo sapiens (man)
C/Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C/Accession: A03903
R/Najarian, R.J.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nes
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A/Title: Primary structure and gene organization of human hepatitis A virus.
A/Reference number: A03903; MUID:85190549; PMID:2986127
A/Accession: A03903
A/Molecule type: Genomic RNA
A/Residues: 1-2227 <NAJ>
A/Cross-references: GB:K02990; NID:G329596; PID:AAA45472.1; PID:G329597
C/Superfamily: hepatitis A virus genome polyprotein
C/Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltran
F/1-245/Product: coat protein 1A #status predicted <CIA>
F/246-491/Product: coat protein 1B #status predicted <CIB>
F/492-836/Product: coat protein 1C #status predicted <CIC>
F/837-980/Product: core protein 2A #status predicted <C2A>
F/981-1076/Product: core protein 2B #status predicted <C2B>
F/1077-1422/Product: core protein 2C #status predicted <C2C>
F/1423-1484/Product: protein 3A #status predicted <C3A>
F/1485-1507/Product: protein 3B #status predicted <C3B>
F/1508-1578/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F/1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGVKQR 20

Db 792 SHIECRPKYKELRLEVGVKQR 811

RESULT 7

GNMYK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; core protein 2C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: A94149, A25914, A94508
R:Cohen, J.L.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with reference sequence
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <COH>
A:Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>
Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
DB 792 SHIECRKPKYKELRLEVGKQR 811

RESULT 8
GNMYB
genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; core protein 2B; core protein 2C; core protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, E.; Deinhardt, V.
Virus Res. 8, 153-171, 1987
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolated from a patient with acute hepatitis A)
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M2073
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-491/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>
Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
DB 792 SHIECRKPKYKELRLEVGKQR 811

RESULT 9
PQ0427
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wide
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Comment: This protein is from the VP1/2A Junction region.
C:Species: human hepatitis A virus genome polyprotein
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 98.1%; Score 105; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 8.3e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
DB 29 SHIECRKPKYKELRLEVGKQR 48

RESULT 10
PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wide
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0429; MUID:92300330; PMID:1318940
A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 97.2%; Score 104; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.2e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
DB 29 SHIECRKPKYKELRLEVGKQR 48

RESULT 11
JH0135
genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994
C:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
Virus Res. 8, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus.
A:Reference number: JH0135; MUID:88129044; PMID:2829458
A:Accession: JH0135
A:Molecule type: genomic RNA

```
C.C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C.Accession: PQ0428
A:Accession: B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nairan, O.V.; Siegl, G.; Wide
J.J. Gen. Virol. 73, 1365-1377, 1992
A.A.Title: Genetic relatedness of hepatitis A virus strains recovered from different geog
A.Reference number: PQ0427; MUID:92300330; PMID:1318940
C.Accession: PQ0428
A:Molecule type: mRNA
A.Residues: 1-56 <ROB>
A.Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypotein

Query Match      89.7%; Score 96; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 2.3e+08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SHIECRPKYKELRLEVQKQR 20
          |||||
Ddb     29 SHIEYRKPYKELRLEVQKQR 48

RESULT 15
PQ0434
genome polyprotein - human hepatitis A virus (strain KPH) (fragment)
C.Species: human hepatitis A virus
C.Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C.Accession: PQ0434
A:Accession: B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nairan, O.V.; Siegl, G.; Wide
J.J. Gen. Virol. 73, 1365-1377, 1992
A.A.Title: Genetic relatedness of hepatitis A virus strains recovered from different geog
A.Reference number: PQ0427; MUID:92300330; PMID:1318940
C.Accession: PQ0434
A:Molecule type: mRNA
A.Residues: 1-56 <ROB>
A.Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polypotein

Query Match      88.8%; Score 95; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 3.3e+08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 SHIECRPKYKELRLEVQKQR 20
          |||||
Ddb     29 SHIEYRKPYKELRLEVQKQR 48

Search completed: March 15, 2004, 14:06:57
Job time : 11 secs
```


AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment)
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RA Lineneyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RA Young A., Mitra S.W.,
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA."
RL J. Virol. 54:247-255(1985).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M10033; AAA5470.1; --
DR PIR; A03904; GNNYHA.
DR InterPro; IPR008975; Viral cap coat.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 837 >852 CORE PROTEIN P2A.
FT NON TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

Query Match 100.0%; Score 107; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 3
POLG HPAV2 STANDARD; PRT; 2226 AA.
ID POLG HPAV2
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic

RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59810; AAA45468.1; --
DR MEROPS; C03.005; --
DR InterPro; IPR004004; Calici.pol.hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR006005; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P3D.
DR InterPro; IPR008975; Viral cap coat.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS; PR00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSN.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 4.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 4
POLG HPAV4
ID POLG HPAV4 STANDARD; PRT; 2226 AA.
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 43C).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean T., Jansen R.W.;

RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59809; AAA45469.1; -;
 DR MEROPS; C03.005; -;
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006055; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;
 Query Match 100.0%; Score 107; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 4.4e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SHIECKPKYKELRLEVGKQR 20
 Db 792 SHIECKPKYKELRLEVGKQR 811
 RESULT 5
 ID POLG_HPAV8 STANDARD; PRT; 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain 18f).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12096;
 RN [1]_TaxID=12096;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RT "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59808; AAA45467.1; -;
 DR PDB; 1QA7; 15-MAY-00.
 DR MEROPS; C03.005; -;
 DR InterPro; IPR004004; Calici_pol_hel.
 DR InterPro; IPR009003; Cys Ser trypsin.
 DR InterPro; IPR006055; RNA_helicase.
 DR InterPro; IPR007095; RNA_pol_DS_PS.
 DR InterPro; IPR001205; RNA_pol_P3D.
 DR InterPro; IPR007094; RNA_pol_PSVir.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR PRINTS; PR00918; CALICIVIRUSNS.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 794
 FT CHAIN 795 900
 FT CHAIN 901 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1495
 FT CHAIN 1496 1518
 FT CHAIN 1519 1737
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251292 MW; 24964A6339C8D6B CRC64;
 Query Match 100.0%; Score 107; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 4.4e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 SHIECKPKYKELRLEVGKQR 20
 Db 792 SHIECKPKYKELRLEVGKQR 811
 RESULT 6
 ID POLG_HPAVH STANDARD; PRT; 2227 AA.
 AC P08617; P06443; Q81082;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain HM-175).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12098;

[1] SEQUENCE FROM N.A.
RN STRAIN=Wild type; PubMed=3023706;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus;
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses";
RL J. Virol. 61:50-59 (1987).
[2] SEQUENCE FROM N.A.
RN STRAIN=Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus;
RT comparison with wild-type virus";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501 (1987).
[3] SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RN MEDLINE=85166289; PubMed=2984684;
RX Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147 (1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M14114; AAA45475.1; -;
CC EMBL: M14707; AAA45465.1; -;
CC EMBL: M14707; AAA45466.1; ALT_INIT.
CC EMBL: M16632; AAA45471.1; -;
CC PIR: A03905; A03905.
CC PIR: A25981; GNNYHM.
CC PIR: A94149; GNNYMK.
CC PDB: 1HAV; 23-DEC-96.
CC MEROPS: C03.005; -;
CC InterPro: IPR004004; Calici_pol_hel.
CC InterPro: IPR009003; Cys_Ser_trypsin.
CC InterPro: IPR000605; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_Pswir.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC PRINTS: PR00318; CALICIVIRUSN.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
KW CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN

DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 432 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 437 980 CORE PROTEIN P2A.
FT CHAIN 981 1076 CORE PROTEIN P2B.
FT CHAIN 1077 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1484 PROBABLE PROTEIN P3A.
FT CHAIN 1485 1507 PROBABLE PROTEIN P3B.
FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVGKQR 20
DB 792 SHIECKPKYKELRLEVGKQR 811

RESULT 8
POLG HPVAVG STANDARD; PRT; 2227 AA.
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
AC Q81090; Q81091; Q81092; Q81093;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain MB).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12100;
RN [1]_TaxID=12100;
RP SEQUENCE FROM N.A.
RX MEDLINE=88045071; PubMed=2823500;
RA Paul A.V., Tada H., der Helm K., Wessel T., Kiehn R., Wimmer E.,
Reinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MB).";
RL Virus Res. 8:153-171(1987).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
{RNA}(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; M20273; AAA45474.1; -.
DR MEROPS; C03.005; -.

DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_ser_trypsin.
DR InterPro; IPR00605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00680; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 432 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 437 980 CORE PROTEIN P2A.
FT CHAIN 981 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 107; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVGKQR 20
DB 792 SHIECKPKYKELRLEVGKQR 811

RESULT 9
POLG HPVAVG STANDARD; PRT; 808 AA.
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein
P2A] (Fragment).
OS Hepatitis A virus (strain GA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31706;
RN [1]_TaxID=31706;
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; M66695; AAA45477.1; -.
DR InterPro; IPR008975; Viral_cap_coat.
KW Polypeptide; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN <1 2 COAT PROTEIN VP4 (P1A).
FT CHAIN 3 223 COAT PROTEIN VP2 (P1B).
FT CHAIN 224 470 COAT PROTEIN VP3 (P1C).

FT CHAIN 471 770 COAT PROTEIN VP1 (P1D).
FT CHAIN 771 >808 CORE PROTEIN P2A.
FT NON_TER 808 808
SQ SEQUENCE 808 AA; 90632 MW; D80CE7B57A479C12 CRC64;

Query Match 88.8%; Score 95; DB 1; Length 808;
Best Local Similarity 95.0%; Pred. No. 1.4e-07; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 1;

QY 1 SHIECKPKYKELRLEVQKQ 20
DB 771 SHIECKPKYKELRLEVQKQ 790

RESULT 10
POLG HPAAVS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12102;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R., Purcell R.H.;
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjapardize A.G., Tsarev S.A., Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D03924; BAA00766.1; --
CC EMBL; X15461; CAA33490.1; --
CC PIR; A30470; GNNYS2.
CC MEROPS; C03.005; --
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR009975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol_1.
CC PFam; PF00910; RNA_helicase; 1.

DR PRINTS; PR00918; CALICVIRUSNS.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 80.4%; Score 86; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 1.1e-05; Indels 0; Gaps 0;
Matches 17; Conservative 1; Mismatches 2;

QY 1 SHIECKPKYKELRLEVQKQ 20
DB 796 SHIECKPKYKELRLEVQKQ 815

RESULT 11
POLG HPAAVT STANDARD; PRT; 839 AA.
AC P31786;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M59286; AAA45473.1; --
CC PIR; J01180; GNNYS2.
CC InterPro; IPR009975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 ?
FT CHAIN ? >839
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CAC04BD1E192DEC CRC64;

Query Match 79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 7.3e-06; Indels 1; Gaps 1;
Matches 18; Conservative 1; Mismatches 0;

OY 1 SHIECKPKYKLEULEVGKOR 20
|||||:|||||:|||||
DB 792 SHIE-KKEYKLEULEVGKOR 810

RESULT 12
YAGA_SCHPO STANDARD; PRT; 420 AA.
AC Q09873;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C12G12.10 in chromosome I.
GN SPAC12G12.10.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21848401; PubMed=11859160;
RX STRAIN=972;
RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,
RA Welljens I., Vanstraelen E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; Z66568; CAA91505.1; -;
DR PIR; S62541; S62541.
DR GenDB_SPombe; SPAC12G12.10; -;
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 1.
DR SMART; SM00320; WD40; 3.
KW Hypothetical protein.
SQ SEQUENCE 420 AA; 47525 MW; 9DC1F710FEB9761 CRC64;
Query Match 43.0%; Score 46; DB 1; Length 420;
Best Local Similarity 46.7%; Pred. No. 6.5;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 4 BCKPKYKLEULEVGK 18

DB 376 DCSLPFKEMRVDDGK 390
:|:|:|:|:|:|
RESULT 13
GCP6_HUMAN STANDARD; PRT; 1819 AA.
ID GCP6_HUMAN
AC Q96P77; Q9BY91; Q9UGX3; Q9UGX4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Gamma-tubulin complex component 6 (GCP-6).
GN TUBGCP6 OR GCP6 OR KIAA1669.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RX MEDLINE=21551508; PubMed=11694571;
RA Murphy S.M., Preble A.M., Patel U.K., O'Connell K.L., Dias D.P.,
RA Moritz M., Agar D., Stults J.T., Stearns T.;
RT "GCP5 and GCP6: two new members of the human gamma-tubulin complex.";
RL Mol. Biol. Cell 12:3340-3352(2001).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Clump M., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Alameda J.P., Babbage A.K.,
RA Baird G.K., Langford C.F., Leversha M., Lloyd C., Lloyd D.M.,
RA Laird I.D., Washreghi-Mohammadi M., Matthews L.H., McLean O.T.,
RA McLay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.B., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Wiley D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoashima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Korf I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
RA Peyrard M., Kedra D., Seroussi E., Franson I., Tapia I., Bruder C.B.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
[3]
RN SEQUENCE OF 33-1819 FROM N.A. (ISOFORM 2).
RP TISSUE=Brain;

```
RX MEDLINE=21156230; PubMed=11258795;
RA Hirosewa M., Nagase T., Murahehi Y., Kikuno R., Ohara O.;
RT "Identification of novel transcribed sequences on human chromosome 22
RL DNA Res. 8:1-9(2001).
CC -!- FUNCTION: Gamma-tubulin complex is necessary for microtubule
CC nucleation at the centrosome.
CC -!- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
CC GCP3, GCP4, GCP5 and GCP6.
CC -!- SUBCELLULAR LOCATION: Centrosome.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q96RT7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q96RT7-2; Sequence=VSP_001624;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the GCP family.
CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
CC in positions 1371 and 1758.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF272887; AAK82968.1; -
DR EMBL; AL022328; CAB63046.1; ALT_SEQ.
DR EMBL; AL022328; CAB63047.1; ALT_SEQ.
DR EMBL; AB051456; BAB33339.1; ALT_FRAME.
DR Genew; HGNC:18127; TUBGCP6.
DR GO; GO:0008274; C:gamma-tubulin ring complex; IDA.
DR GO; GO:0008017; F:microtubule binding; IDA.
DR GO; GO:0007020; F:microtubule nucleation; IDA.
DR InterPro; IPR007259; SPC97_Spc98.
DR Pfam; PF04130; SPC97_Spc98; 1.
DR Microtubule; Repeat; Alternative splicing.
DR DOMAIN; 1027 1269 9 X 27 AA TANDEM REPEATS.
FT REPEAT 1027 1053 1.
FT REPEAT 1054 1080 2.
FT REPEAT 1081 1107 3.
FT REPEAT 1108 1134 4.
FT REPEAT 1135 1161 5.
FT REPEAT 1162 1188 6.
FT REPEAT 1189 1215 7.
FT REPEAT 1216 1242 8.
FT REPEAT 1243 1269 9.
FT VARSPLIC 1724 1757 Missing (in isoform 2).
FT FTID=VSP_001624.
FT CONFLICT 567 567 S -> L (IN REF. 3).
FT CONFLICT 1377 1377 A -> T (IN REF. 2 AND 3).
FT CONFLICT 1621 1621 L -> V (IN REF. 2).
SQ SEQUENCE 1819 AA; 200455 MW; 923576544D34594A CRC64;

Query Match 42.1%; Score 45; DB 1; Length 1819;
Best Local Similarity 50.0%; Pred. NO. 41;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 HIECRPKYKELRLEVGKQ 19
Db 653 HSSVSKEEKLRLMEIAKQ 670

RESULT 14
CO3_RAT STANDARD; PRT; 1663 AA.
AC P01026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
```

```
DB GN
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Liver;
RX MEDLINE=90245672; PubMed=2336397;
RA Misumi Y., Sohda M., Ikehara Y.;
RT "Nucleotide and deduced amino acid sequence of rat complement C3.";
RL Nucleic Acids Res. 18:2178-2178(1990).
RN (2)
RP SEQUENCE OF 571-748.
RX MEDLINE=7906262; PubMed=309768;
RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA Daniels J.S., Daughaday W.H., Bradshaw R.A.;
RT "Purification, characterization, and amino acid sequence of rat
RT anaphylatoxin (C3a)."
RL Biochemistry 17:5031-5038(1978).
RN (3)
RP SEQUENCE OF 1316-1595 FROM N.A.
RX MEDLINE=89380332; PubMed=2674144;
RA Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA Lyttle C.R.;
RT "Estrogen regulation of tissue-specific expression of complement C3.";
RL J. Biol. Chem. 264:16941-16947(1989).
CC -!- FUNCTION: C3 plays a central role in the activation of the
CC complement system. Its processing by C3 convertase is the central
CC reaction in both classical and alternative complement pathways.
CC After activation C3b can bind covalently, via its reactive
CC thioester, to cell surface carbohydrates or immune aggregates.
CC -!- FUNCTION: Derived from proteolytic degradation of complement C3,
CC C3a anaphylatoxin is a mediator of local inflammatory process. It
CC induces the contraction of smooth muscle, increases vascular
CC permeability and causes histamine release from mast cells and
CC basophilic leukocytes.
CC -!- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG
CC RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE
CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,
CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA
CC CHAIN).
CC -!- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC -!- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC -!- SIMILARITY: Contains 1 NTR domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X52477; CAA36716.1; -
DR EMBL; M29866; AAA40837.1; ALT_SEQ.
DR PIR; S15764; C3ET.
DR PDB; 1QQF; 31-JUL-00.
DR PDB; 1QSU; 31-JUL-00.
DR InterPro; IPR002890; A2M_N.
DR InterPro; IPR009048; AM_receptor_bind.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR001840; Anaphylatoxn.
DR InterPro; IPR001599; MacrogloblnA2.
DR InterPro; IPR001134; Netrin_C.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR Pfam; PF01821; ANATO; 1.
DR Pfam; PF01759; NTR; 1.
DR PRINTS; PR00004; ANAPHYLATOXN.
DR ProDom; PD003264; Anaphylatoxin; 1.
```

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; D00736; BAA00637.1; -

DR PIR; C32063; VHN23

DR InterPro; IPR004930; Pneumo_ncap.

DR Pfam; PF03246; Pneumo_ncap_1.

DR ProDom; PD006438; Pneumo_ncap; 1.

KW Nucleocapsid.

SQ SEQUENCE 391 AA; 43416 MW; 9A2BBB50103B2835 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 391;

Best Local Similarity 56.2%; Pred. No. 18;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 IECKRPYKELRLEVGK 18

DB 129 IESRRKSYKKLKEMGE 144

Search completed: March 15, 2004, 14:00:57

Job time : 6.64706 secs

SWART; SM00104; ANATO; 1.

SWART; SM00643; C345C; 1.

DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.

DR PROSITE; PS01177; ANAPHYLATOXIN_1; 1.

DR PROSITE; PS01178; ANAPHYLATOXIN_2; 1.

DR PROSITE; PS0189; NTR; 1.

DR Complement pathway; Complement alternate pathway; Plasma; Inflammatory response; Glycoprotein; Signal; 3D-structure; Thioester bond.

KW SIGNAL 1 24

FT CHAIN 25 1663 COMPLEMENT C3.

FT CHAIN 25 666 BETA CHAIN.

FT CHAIN 671 1663 ALPHA CHAIN.

FT PEPTIDE 671 748 C3A ANAPHYLATOXIN.

FT CHAIN 749 1663 C3B (ALPHA' CHAIN).

FT DOMAIN 693 728 ANAPHYLATOXIN-LIKE.

FT DOMAIN 1518 1661 NTR.

FT SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).

FT DISULFID 558 816 INTERCHAIN (BY SIMILARITY).

FT DISULFID 626 661 BY SIMILARITY.

FT DISULFID 693 720 BY SIMILARITY.

FT DISULFID 694 727 BY SIMILARITY.

FT DISULFID 707 728 BY SIMILARITY.

FT DISULFID 873 1513 BY SIMILARITY.

FT DISULFID 1101 1138 BY SIMILARITY.

FT DISULFID 1358 1489 BY SIMILARITY.

FT DISULFID 1389 1458 BY SIMILARITY.

FT DISULFID 1506 1511 BY SIMILARITY.

FT DISULFID 1518 1590 BY SIMILARITY.

FT DISULFID 1537 1661 BY SIMILARITY.

FT CROSSLNK 1010 1013 Isoglutamyl cysteine thioester (Cys-Gln).

FT CARBOHYD 939 939 N-LINKED (GLCNAC. .) (PROBABLE).

FT CARBOHYD 1617 1617 N-LINKED (GLCNAC. .) (PROBABLE).

FT CONFLICT 721 722 LK -> KL (IN REF. 2).

SQ SEQUENCE 1663 AA; 186460 MW; 2F87CCB143CDD4BC CRC64;

Query Match 40.7%; Score 43.5; DB 1; Length 1663;

Best Local Similarity 47.4%; Pred. No. 65;

Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

QY 1 SHIECKRPYKELRLEVGKQ 19

DB 1586 SHVKCR---NALKLGKQKQ 1601

RESULT 15

NCAP_HRSV1 STANDARD; PRT; 391 AA.

ID NCAP_HRSV1 AC F24566;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 01-MAR-1992 (Rel. 21, Last annotation update)

DE Nucleocapsid protein.

GN N.

OS Human respiratory syncytial virus (subgroup B / strain 18537).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI_TaxID=11251;

RN [1]

SEQUENCE FROM N.A.

RP MEDLINE=89279331; PubMed=2525176;

RX Johnson P.R., Collins P.L.;

RA "The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (RSV) of antigenic subgroups A and B: sequence conservation and divergence within RSV genomic RNA."

RT J. Gen. Virol. 70:1539-1547(1989).

RL J. Gen. Virol. 70:1539-1547(1989).

CC -!- FUNCTION: Most abundant protein in the virion and an important element conferring helical symmetry on the nucleoprotein core as well as interacting with m protein during virion formation.

CC -!- SIMILARITY: Belongs to the paramyxoviruses nucleocapsid family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:24:17 ; Search time 44.2353 Seconds
(without alignments)
127.748 Million cell updates/sec

Title: US-09-171-432A-39
Perfect score: 107
Sequence: 1 SHIECKPKYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Genesecp1980s:*
2: Genesecp1990s:*
3: Genesecp2000s:*
4: Genesecp2001s:*
5: Genesecp2002s:*
6: Genesecp2003as:*
7: Genesecp2003bs:*
8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
1	107	100.0	20	2	AAW42922 Immunogen
2	107	100.0	21	4	AAB69439 Synthetic
3	107	100.0	366	1	AAP50230 Sequence
4	107	100.0	854	1	AAP50287 Hepatitis
5	107	100.0	993	1	AAP50116 Sequence
6	107	100.0	993	1	AAP50231 Sequence
7	107	100.0	1077	2	AAW95559 A partial
8	107	100.0	1091	2	AAW32426 Translate
9	107	100.0	2227	2	AAW5697 Attenuate
10	107	100.0	2227	2	AAW34074 Hepatitis
11	107	100.0	2227	3	AAB18609 Amino aci
12	107	100.0	2227	3	AAB18607 Amino aci
13	107	100.0	2227	3	AAB18608 Amino aci
14	107	100.0	2227	5	AAE19899 Hepatitis
15	107	100.0	2227	5	ABG31729 Attenuate
16	107	100.0	2227	5	ABG31727 Wild-type
17	107	100.0	2227	5	ABG31728 Hepatitis
18	107	100.0	2227	6	ABU08640 Attenuat
19	107	100.0	2227	6	ABU08641 Attenuat
20	107	100.0	2227	6	ABU08639 Wild type
21	107	100.0	2227	6	ABU08635 Hepatitis
22	91	85.0	2227	1	ABW00350 Sequence
23	84.5	79.0	839	2	AAE15629 Capsid re
24	67	62.6	20	2	AAW42923 Immunogen
25	67	62.6	21	4	AAB69440 Synthetic

26	45.5	42.5	178	4	AAU03662	Aau03662 Group B S
27	45.5	42.5	330	5	ABP30218	Abp30218 Streptoco
28	45.5	42.5	341	5	ABP25499	Abp25499 Streptoco
29	45	42.1	1273	5	ABP62888	Abp62888 Human pol
30	45	42.1	1437	6	ABP98857	Abp98857 Human str
31	45	42.1	1819	7	AD660309	Ad660309 Human Pro
32	44	41.1	71	3	AAG07411	Aag07411 Arabidops
33	44	41.1	88	3	AAE13254	Aae13254 Arabidops
34	44	41.1	211	1	AAE13255	Aae13255 Arabidops
35	44	41.1	346	2	AAE13255	Aae13255 Human tum
36	43	40.2	45	4	AAE13255	Aae13255 Human tum
37	43	40.2	45	5	AAE13255	Aae13255 Human tum
38	43	40.2	52	4	AAE13254	Aae13254 Human tum
39	43	40.2	52	5	AAE13254	Aae13254 Human tum
40	43	40.2	141	7	AD660309	Ad660309 Human Pro
41	43	40.2	186	6	ABM69585	Abm69585 Photornab
42	43	40.2	287	3	AAE13255	Aae13255 Arabidops
43	43	40.2	312	3	AAE13255	Aae13255 Arabidops
44	43	40.2	320	3	AAE13255	Aae13255 Arabidops
45	43	40.2	332	5	ABE91335	Abb91335 Herbitida

ALIGNMENTS

RESULT 1
AAW42922
ID AAW42922 standard; peptide; 20 AA.
XX AC AAW42922;
XX DT 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1315.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX OS Synthetic.
OS Hepatitis A virus.
XX PN WO9740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
(USSH) US DEPT HEALTH & HUMAN SERVICES.
XX Fields HA, Khudyakov YE;
XX WPI; 1997-535831/49.
Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of antibodies against HAV in a mammal.

Claim 18; Page 112; 14Opp; English.
Peptides AAW42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 792-811. The present peptide is derived from amino acids 792-811, and has a reactivity of 54.2% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The peptides can also be used to make an antibody against HAV by administering the peptide to a mammal

Query Match 100.0%; Score 107; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPYKELRLEVGKQR 20
DB 1 SHIECRPYKELRLEVGKQR 20

RESULT 2
AAB69439
ID AAB69439 standard; peptide; 21 AA.
AC AAB69439;
XX
XX 20-APR-2001 (first entry)
XX
XX Synthetic HAV P2A peptide, SEQ ID NO: 39.
XX
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
KW
XX
XX Hepatitis A virus.
OS Synthetic.
OS
XX WO200105924-A2.
PN
XX
XX 25-JAN-2001.
XX
XX 14-JUL-2000; 2000WO-US019267.
PF
XX 15-JUL-1999; 99US-0144412P.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Fielde HA, Khudyakov YE;
XX
XX WPI; 2001-112681/12.
DR
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
PT
XX Claim 13; Page 93; 130pp; English.
XX

The present sequence is one of a number of synthetic peptides which are immunoreactive with hepatitis A virus (HAV) antibodies. The peptides comprise antigenic epitopes of the major structural capsid polypeptides or non-structural polypeptides of HAV with one or more glutamine molecules at the carboxy end of the peptide. The peptides are used to detect the presence of antibodies against HAV in mammalian serum, to detect the presence of HAV in a human or animal through the binding of the peptide to an antibody, to detect acute phase infection by detecting IGM antibodies in mammalian serum and detecting convalescence in a mammal. The peptides are used to detect or quantify HAV antibodies in samples in clinical or research-based assays using immunoblotting, fluorescence in situ hybridisation analysis, gel-mobility shift assays, tracking of radioactive or bioluminescent markers, chromatography or electrophoresis. The peptides are used to induce an immune response to HAV when administered to a human or animal. Glutamine at the carboxy end of the peptides enhances the IGM antibody reactivity

Sequence 21 AA;
Query Match 100.0%; Score 107; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPYKELRLEVGKQR 20
DB 1 SHIECRPYKELRLEVGKQR 20

RESULT 3
AAP50230
ID AAP50230 standard; protein; 366 AA.
XX
XX AAP50230;
AC
XX 28-NOV-1991 (first entry)
DT
XX Sequence of hepatitis A virus (HAV) surface protein (VP-1).
DE
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
KW
XX Hepatitis A virus.
OS
XX EPI38704-A.
PN
XX 24-APR-1985.
PD
XX 09-OCT-1984; 84EP-00402025.
PF
XX 14-OCT-1983; 83US-00541836.
PR
XX 02-MAR-1984; 84US-00585942.
XX
XX (MERI) MERCK & CO INC.
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
PI
XX WPI; 1985-100819/17.
DR
XX N-ESDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus.
PT
XX Claim 21; Page 46-48; 49pp; English.
PS
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 daltons is sepd
CC
XX
SQ Sequence 366 AA;
Query Match 100.0%; Score 107; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 3.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPYKELRLEVGKQR 20
DB 301 SHIECRPYKELRLEVGKQR 320

RESULT 4
AAP50287
ID AAP50287 standard; protein; 854 AA.
XX
XX AAP50287;
AC
XX 25-MAR-2003 (revised)
DT
XX 30-NOV-1991 (first entry)
DT
XX Hepatitis A virus (HAV) peptide corresponding to the capsid protein
DE region of poliovirus RNA.
DE
XX Hepatitis A virus assay; antigen; antibody.
KW
XX Hepatitis A virus.
OS
XX WO8501517-A.
PN
XX 11-APR-1985.
PD
XX 27-SEP-1984; 84WO-US001552.
PF
XX

PR 30-SEP-1983; 83US-00537911.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PA Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racanello VR;
 XX WPI; 1985-098846/16.
 DR N-PSDB; AAN50330.
 XX New hepatitis A virus CDNA - useful in assays for the virus and for
 PT prodn. of the viral antigen and antibodies to it.
 XX Example; Fig 7; 60pp; English.
 XX The inventors claim HAV cDNA and a method for producing it, whereby large
 CC amts. can be obt'd. economically. The cDNA is useful in the assay for
 CC detection of HAV quickly and easily and with high sensitivity and
 CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
 CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX
 SQ Sequence 854 AA;
 Query Match 100.0%; Score 107; DB 1; Length 854;
 Best Local Similarity 100.0%; Pred. No. 8e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIECKPKYKELRLEVQKOR 20
 DB 792 SHIECKPKYKELRLEVQKOR 811
 RESULT 5
 AAP50116
 ID AAP50116 standard; protein; 993 AA.
 XX
 AC AAP50116;
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 30-SEP-1991 (first entry)
 XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
 DE and VP-4.
 XX Antigenic protein; immunogen; vaccine.
 XX Hepatitis A virus; (strain CR326).
 XX EP154587-A.
 XX 11-SEP-1985.
 XX 27-FEB-1985; 85EP-00400369.
 XX 02-MAR-1984; 84US-00585819.
 XX (MERI) MERCK & CO INC.
 XX Linemeyer DL, Menke JG, Rueben RG, Mitra SW;
 XX WPI; 1985-224964/37.
 DR N-PSDB; AAN50139.
 XX New nucleotide sequences coding for hepatitis A virus antigens - useful
 PT for eliciting normal immune response and in vaccines for protecting
 PT against the virus.
 XX Example; Page 11-17; 32pp; English.
 XX Within the sequence in AAN50139 is encoded the information necessary to
 CC make the antigenic proteins of HAV. The sequences encoding for the

CC structural proteins begin at base 403. The key sub-unit sequences within
 CC VP-1, designated Sequences I, II, III, IV, and V, start, respectively at
 CC 1882, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
 CC valuable as encoding antigenic proteins are the sequences from base 1749
 CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
 CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
 CC vector for producing antigen protein. Sequences II-V are claimed. X in
 CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
 CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX
 SQ Sequence 993 AA;
 Query Match 100.0%; Score 107; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 9.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SHIECKPKYKELRLEVQKOR 20
 DB 928 SHIECKPKYKELRLEVQKOR 947
 RESULT 6
 AAP50231
 ID AAP50231 standard; protein; 993 AA.
 XX
 AC AAP50231;
 DT 28-NOV-1991 (first entry)
 XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
 DE including surface protein (VP-1).
 XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 XX Hepatitis A virus.
 XX Key Location/Qualifiers
 FT Protein 628..993
 FT /note="claimed; X denotes translated stop codons and
 FT unspecified triplets"
 XX EP138704-A.
 XX 24-APR-1985.
 XX 09-OCT-1984; 84EP-00402025.
 XX 14-OCT-1983; 83US-00541836.
 PR 02-MAR-1984; 84US-00585942.
 XX (MERI) MERCK & CO INC.
 XX Hughes JV, Scolnick EM, Tomassini JF;
 PI WPI; 1985-100818/17.
 DR N-PSDB; AAN50274.
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus.
 XX Disclosure; Page 17-23; 49pp; English.
 XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
 CC surfactant and a reducing agent. The viral proteins are sepd. and the
 CC protein of molecular wt. 33000 daltons is sepd
 XX Sequence 993 AA;
 Query Match 100.0%; Score 107; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 9.4e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
 Db 928 SHIECRKPKYKELRLEVGKQR 947

RESULT 7
 AAW95559
 ID AAW95559 standard; protein; 1077 AA.
 XX AC AAW95559;
 XX DT 28-APR-1999 (first entry)
 XX DE A partial hepatitis A virus (HAV) protein.
 XX Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX OS Hepatitis A virus.
 XX PN US5849562-A.
 XX PD 15-DEC-1998.
 XX PF 06-JUN-1995; 95US-00468926.
 XX PR 30-SEP-1983; 83US-00537911.
 XX PR 27-SEP-1984; 84US-00654942.
 XX PR 06-OCT-1988; 88US-00256135.
 XX PR 06-NOV-1991; 91US-00788262.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Emerson SU, Purcell RH;
 XX WPI; 1993-094412/08.
 XX DR N-PSDB; AAX01006.
 XX Chimeric hepatitis A virus strains - with P2 region from cell-culture-
 PT adapted strain in wild-type genome.
 XX PS Disclosure; Fig 7A-L; 36pp; English.
 CC The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting of a wild
 CC -type HAV genome in which the P2 region is replaced by the P2 region from
 CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
 CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
 CC sufficient for establishment of infection and accelerated growth in cell
 CC culture
 XX SQ Sequence 1077 AA;
 Query Match 100.0%; Score 107; DB 2; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
 Db 1015 SHIECRKPKYKELRLEVGKQR 1034

RESULT 8
 AAR32426
 ID AAR32426 standard; protein; 1091 AA.
 XX AC AAR32426;
 XX DT 27-AUG-2003 (revised)
 XX DT 25-MAR-2003 (revised)
 XX DT 17-DEC-2001 (revised)
 XX DT 10-JUN-1993 (first entry)

DE Translated from 5' region of Hepatitis A Virus genomic clone.
 KW HAV HM-175; chronic liver disease; picornavirus.
 XX Hepatitis A virus.
 OS
 XX Key Location/Qualifiers
 FH 1..711
 FT /note= "X's correspond to nonsense codons, i.e. this
 FT region is not an ORF"
 FT 238..1091
 FT /label= ORF
 FT /note= "second putative initiation codon at position 240"
 XX USN788262-N.
 XX 15-DEC-1992.
 XX 06-NOV-1991; 91US-00788262.
 XX 30-SEP-1983; 88US-00536911.
 XX 27-SEP-1984; 84US-00654942.
 XX 06-OCT-1988; 88US-00256135.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX Ticehurst JR, Baltimore D, Feinstone SM, Purcell RH;
 XX Racaniello VR, Baroudy BM, Emerson SU;
 XX WPI; 1993-067429/08.
 XX DR N-PSDB; AAQ36934.
 XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
 PT antigen and antibodies.
 XX PS Disclosure; Fig 7; 65pp; English.
 XX HAV virion RNA was extracted from the livers of marmosets which had been
 CC inoculated with HAV (the HAV had previously been passaged twice in
 CC marmosets). The RNA was used to prepare ds cDNA clones by standard
 CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected
 CC African Green Monkey Kidney cells were selected for further analysis. A
 CC 7.4kb restriction map (about 98% of the HAV genome) was constructed from
 CC 5 overlapping inserts. The sequence of the first 3.3kb (approx.) from the
 CC 5'-terminus was determined. An amino acid sequence was deduced from the
 CC entire clone and an open reading frame was identified starting at
 CC position 238. A comparison of the predicted HAV amino acid sequences with
 CC the known capsid protein sequences of other picornaviruses (poliovirus,
 CC foot and mouth disease virus and encephalomyelitis virus) revealed areas
 CC of local homology. (Note: Revised entry submitted to correct the patent
 CC number format of US Government-owned NPL applications to prevent clashes
 CC with ongoing US granted patent numbers. For further information please
 CC visit the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html.)
 CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX SQ Sequence 1091 AA;
 Query Match 100.0%; Score 107; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20
 Db 1029 SHIECRKPKYKELRLEVGKQR 1048

RESULT 9
 AAR05697
 ID AAR05697 standard; protein; 2227 AA.
 XX AC AAR05697;
 XX

DT 24-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 15-AUG-1990 (first entry)
XX
DE Attenuated hepatitis A virus.
XX
KW Hepatitis A virus; vaccine; attenuated.
XX
OS Hepatitis A virus; strain HM-175.
XX
FH Key Location/Qualifiers
FT Region 1. .23
FT /label= VP4 = 1A
FT Region 24. .245
FT /label= VP2 = 1B
FT Region 246. .491
FT /label= VP3 = 1C
FT Region 492. .791
FT /label= VP1 = 1D
FT Region 792. .980
FT /label= 2A
FT Region 981. .1087
FT /label= 2B
FT Region 1088. .1422
FT /label= 2C
FT Region 1423. .1496
FT /label= 3A
FT Region 1497. .1519
FT /label= 3B = VPg
FT Region 1520. .1738
FT /label= 3C
FT Region 1739. .2227
FT /label= 3D
XX
PN US48994228-A.
XX
PD 16-JAN-1990.
XX
XX 12-JUL-1988; 88US-00217824.
XX
PR 19-SEP-1984; 84US-00652067.
PR 09-SEP-1986; 86US-00905146.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
XX
PI Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
PI Daemer RJ, Gust ID;
XX
DR WPI; 1990-075557/10.
DR N-PSDB; AAQ03512.
XX
PT Vaccine against hepatitis A virus infection - comprises novel attenuated
PT hepatitis A virus strain.
XX
PS Claim 1; Fig 1; 18pp; English.
XX
XX The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PT field.)
CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 2; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLVGVKQR 20
DB 792 SHIECKPKYKELRLVGVKQR 811
RESULT 10
AAW34074
ID AAW34074 standard; protein; 2227 AA.
XX
AC AAW34074;
XX
DT 17-OCT-2003 (revised)
DT 27-APR-1998 (first entry)
XX
DE Hepatitis A virus HM-175 protein sequence.
XX
KW HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
KW vaccine.
XX
OS Hepatitis A virus; HM-175.
XX
FH Key Location/Qualifiers
FT Protein 1. .23
FT /label= VP4
FT Protein 24. .245
FT /label= VP2
FT Protein 246. .491
FT /label= VP3
FT Protein 492. .791
FT /label= VP1
FT Protein 792. .980
FT /label= 2A
FT Protein 981. .1087
FT /label= 2B
FT Protein 1088. .1422
FT /label= 2C
FT Protein 1423. .1496
FT /label= 3A
FT Protein 1497. .1519
FT /label= 3B
FT Protein 1520. .1738
FT /label= 3C
FT Protein 1739. .2227
FT /label= 3D
XX
PN WO9740166-A2.
XX
XX 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006506.
XX
PR 19-APR-1996; 96US-0015642P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Raychaudhuri G, Emerson SU, Purcell RH;
PI
DR WPI; 1997-535850/49.
DR N-PSDB; AAT93023.
XX
XX Human attenuated HAV genome containing simian HAV 2C gene - useful as
PT vaccines against HAV infection.
XX
PS Disclosure; Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
CC 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
CC by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
CC construct (I) comprises a genome of HAV, where the genome is a human
CC attenuated HAV genome in which a region of the 2C gene has been replaced
CC (see AAT93024). The region of the 2C gene from AGM-27 contained in the
CC construct preferably encodes amino acids 120-328 of the 2C protein, amino

CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can also be used to stimulate the production of protective antibodies in the mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 2; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

QY 1 SHIECRPKYKELRLEVGKQR 20

|||||

DB 792 SHIECRPKYKELRLEVGKQR 811

RESULT 11

AA18609

ID AAB18609 standard; protein; 2227 AA.

XX AAB18609;

XX 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

KW HAV 4380.

XX Hepatitis A virus.

XX US6113912-A.

XX 05-SEP-2000.

XX 07-JUN-1995; 95US-00475886.

XX 18-SEP-1992; 92US-00947338.

XX 17-SEP-1993; 93WO-US008610.

XX 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

XX N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.

XX Disclosure; Col 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates

XX Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 3; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20

|||||

DB 792 SHIECRPKYKELRLEVGKQR 811

RESULT 12

AA18607

ID AAB18607 standard; protein; 2227 AA.

XX AAB18607;

XX 15-JAN-2001 (first entry)

DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

XX US6113912-A.

XX 05-SEP-2000.

XX 07-JUN-1995; 95US-00475886.

XX 18-SEP-1992; 92US-00947338.

XX 17-SEP-1993; 93WO-US008610.

XX 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

XX N-PSDB; AAA75476.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell line useful as vaccine for protecting humans against hepatitis A virus infection, has modified genome compared to wild type.

XX Disclosure; Fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus (HAV) strain HM-174. The sequence is modified to produce HAV which are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates

XX Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 3; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20

|||||

DB 792 SHIECRPKYKELRLEVGKQR 811

RESULT 13

AA18608

ID AAB18608 standard; protein; 2227 AA.

XX AAB18608;

XX 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

KW P-35 virus.

XX Hepatitis A virus.

OS Hepatitis A virus.

XX US6113912-A.
XX 05-SEP-2000.
XX 07-JUN-1995; 95US-00475886.
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX N-PSDB; AAA75477.
XX WPI; 2000-586464/55.
XX N-PSDB; AAA75477.
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type.
XX Disclosure; Col 67-78; 72pp; English.
XX The present sequence is derived from passage 35 of a wild type hepatitis
XX A virus (HAV) strain HM-174. The resulting virus is designated P-35
XX virus. The sequence is modified to produce HAV which are adapted to
XX growth in the human fibroblast-like cell line MRC-5. The HAV is able to
XX propagate in MRC-5 cells and retain appropriate attenuation. It is useful
XX as a live vaccine for prophylaxis of hepatitis A in humans and other
XX primates.
XX Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECKPKYKELRLEVGKOR 20
DB 792 SHIECKPKYKELRLEVGKOR 811
RESULT 14
AAE19899
ID AAE19899 standard; protein; 2227 AA.
XX AC AAE19899;
XX 18-JUN-2002 (first entry)
XX Hepatitis A virus (HAV) protein.
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX Hepatitis A virus.
XX WO200213855-A2.
XX 21-FEB-2002.
XX 15-AUG-2001; 2001WO-IB001808.
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX 03-NOV-2000; 2000US-00705547.
XX (TRIP-) TRIPEP AB.
XX Sallberg M, Hultgren C;
XX WPI; 2002-241837/29.
XX N-PSDB; AAD31766.

XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus.
XX Claim 11; Page 82-87; 120pp; English.
XX The invention relates to a composition comprising ribavirin and an
XX antigen preferably non structural 3 protein (NS3)/4A fragment of
XX hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX sequence. The composition is useful for enhancing an immune response to a
XX hepatitis C antigen in humans, domestic, sport or pet species and as
XX vaccines for treating and preventing HCV infections. The composition is
XX also useful for treating viral, bacterial, fungal diseases and cancer.
XX The present sequence is hepatitis A virus (HAV) protein
XX Sequence 2227 AA;
Query Match 100.0%; Score 107; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECKPKYKELRLEVGKOR 20
DB 792 SHIECKPKYKELRLEVGKOR 811
RESULT 15
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX AC ABG31729;
XX 29-AUG-2003 (revised)
XX 29-NOV-2002 (first entry)
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX HAV 4380.
XX Hepatitis A virus; strain HM-175.
XX US6423318-B1.
XX 23-JUL-2002.
XX 31-AUG-2000; 2000US-00653499.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX 07-JUN-1995; 95US-00475886.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
XX N-PSDB; AB852789.
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX Disclosure; Col 93-104; 71pp; English.
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents an attenuated
XX hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
XX standardise OS field)

SQ Sequence 2227 AA;

Query Match 100.0%; Score 107; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKQR 20
| | | | | | | | | | | | | | | | | | | | | |
Db 792 SHIECRKPYKELRLEVGKQR 811

Search completed: March 15, 2004, 13:59:56
Job time : 45.2353 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: March 15, 2004, 13:15:47 ; Search time 11.2941 Seconds
(without alignments)
91.421 Million cell updates/sec

Title: US-09-171-432A-39
Perfect score: 107
Sequence: 1 SHIECKPKYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap: *
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap: *
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap: *
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap: *
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap: *
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	1091	6	Patent No. 5516630-2
2	107	100.0	2227	3	US-08-475-886-2
3	107	100.0	2227	3	US-08-475-886-4
4	107	100.0	2227	3	US-08-475-886-6
5	107	100.0	2227	3	US-08-397-232-2
6	107	100.0	2227	3	US-08-397-232-4
7	107	100.0	2227	3	US-09-171-387-2
8	107	100.0	2227	4	US-09-653-499-2
9	107	100.0	2227	4	US-09-653-499-4
10	107	100.0	2227	4	US-09-653-499-6
11	107	100.0	2227	4	US-10-104-968-12
12	107	100.0	2227	4	US-10-135-988-2
13	107	100.0	2227	4	US-10-135-988-4
14	107	100.0	2227	4	US-10-135-988-6
15	84.5	79.0	839	1	US-08-087-016-2
16	43	40.2	1127	3	US-09-150-460B-11
17	42.5	39.7	5405	3	US-08-718-388-9
18	42	39.3	444	4	US-09-252-991A-32415
19	42	39.3	472	4	US-09-328-352-6296
20	42	39.3	977	4	US-09-543-681A-6386
21	41	38.3	114	4	US-09-489-847-321
22	41	38.3	153	4	US-09-533-029-96
23	41	38.3	391	5	PCT-US91-08177-3
24	41	38.3	534	4	US-09-252-991A-32086
25	41	38.3	607	4	US-09-907-794A-190
26	41	38.3	607	4	US-09-905-125A-190
27	41	38.3	607	4	US-09-902-775A-190

28	41	38.3	843	4	US-09-491-356C-20	Sequence 20, Appl
29	41	38.3	4544	1	US-08-469-486-52	Sequence 52, Appl
30	41	38.3	4544	2	US-08-469-658-52	Sequence 52, Appl
31	40	37.4	110	4	US-09-341-461-21	Sequence 21, Appl
32	40	37.4	414	4	US-09-252-991A-26659	Sequence 26659, A
33	39	36.4	24	1	US-08-406-347A-13	Sequence 13, Appl
34	39	36.4	154	4	US-09-198-452A-530	Sequence 530, Appl
35	39	36.4	238	4	US-09-257-179-80	Sequence 80, Appl
36	39	36.4	305	4	US-09-635-877A-3	Sequence 3, Appl
37	39	36.4	305	4	US-09-636-060C-3	Sequence 3, Appl
38	39	36.4	305	4	US-09-636-060C-3	Sequence 3, Appl
39	39	36.4	305	4	US-09-986-552-3	Sequence 3, Appl
40	39	36.4	344	3	US-09-393-554-2	Sequence 2, Appl
41	39	36.4	495	4	US-09-252-991A-25802	Sequence 25802, A
42	39	36.4	520	6	5223391-5	Patent No. 5223391
43	39	36.4	892	4	US-09-585-858-16	Sequence 16, Appl
44	39	36.4	1045	4	US-09-489-039A-8589	Sequence 8589, Ap
45	39	36.4	1124	4	US-09-252-991A-26810	Sequence 26810, A

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICHEURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE, STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.; BAROUDY, BAIGIE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 107; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 8.4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPKYKELRLEVGKQR 20
Db 1029 SHIECKPKYKELRLEVGKQR 1048

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1

```
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match
Best Local Similarity 100.0%; Score 107; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match
Best Local Similarity 100.0%; Score 107; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match
Best Local Similarity 100.0%; Score 107; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match
Best Local Similarity 100.0%; Score 107; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLEVGKQR 20
Db 792 SHIECRPKYKELRLEVGKQR 811

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match
Best Local Similarity 100.0%; Score 107; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



```
QY 1 SHIECKPKYKELRLEVGKQR 20
Db 792 SHIECKPKYKELRLEVGKQR 811

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feiler
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 751-6849
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 107; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVGKQR 20
Db 792 SHIECKPKYKELRLEVGKQR 811

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-2

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVGKQR 20
Db 792 SHIECKPKYKELRLEVGKQR 811

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECKPKYKELRLEVGKQR 20
Db 792 SHIECKPKYKELRLEVGKQR 811

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; EMERSON, SUZANNE U
; PURCELL, ROBERT H
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
```

; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; SOFTWARE: Patent in Ver. 2.1
; NUMBER OF SEQ ID NOS: 6
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2
Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKQR 20
DB 792 SHIECRPKYKELRLEVGKQR 811
RESULT 13
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4
Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKQR 20
DB 792 SHIECRPKYKELRLEVGKQR 811
RESULT 14
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6

; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; SOFTWARE: Patent in Ver. 2.1
; NUMBER OF SEQ ID NOS: 6
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6
Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKQR 20
DB 792 SHIECRPKYKELRLEVGKQR 811
RESULT 11
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRPKYKELRLEVGKQR 20
DB 792 SHIECRPKYKELRLEVGKQR 811
RESULT 12
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18

; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVKGQR 20
DB 792 SHIECRKPYKELRLEVKGQR 811
|||||:|||||

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OVANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 4.4e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 SHIECRKPYKELRLEVKGQR 20
DB 792 SHIE-KPYKELRLEVKGQR 810
|||||:|||||

Search completed: March 15, 2004, 13:26:03
Job time : 12.2941 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:23:07 ; Search time 23.1765 Seconds
(without alignments)
182.213 Million cell updates/sec

Title: US-09-171-432A-39
Perfect score: 107
Sequence: 1 SHIECRKPYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Match Length	Description
1	107	100.0	Sequence 39, Appl
2	107	100.0	Sequence 45, Appl
3	107	100.0	Sequence 40, Appl
4	107	100.0	Sequence 41, Appl
5	107	100.0	Sequence 12, Appl
6	107	100.0	Sequence 12, Appl
7	107	100.0	Sequence 2, Appl
8	107	100.0	Sequence 4, Appl
9	107	100.0	Sequence 6, Appl
10	67	62.6	Sequence 40, Appl
11	45.5	42.5	Sequence 124, Appl
12	43	40.2	Sequence 13, Appl
13	43	40.2	Sequence 12, Appl
14	43	40.2	Sequence 2, Appl
15	42.5	39.7	Sequence 1116, Appl

ALIGNMENTS

RESULT 1
US-09-171-432A-39
; Sequence 39, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

16	42.5	39.7	5405	13	US-10-025-380-1116	Sequence 1116, Ap
17	42	39.3	1087	15	US-10-369-493-6801	Sequence 6801, Ap
18	42	39.3	1289	9	US-09-932-145-11	Sequence 11, Appl
19	42	39.3	1289	15	US-10-395-812-11	Sequence 11, Appl
20	41	38.3	22	14	US-10-173-461-21	Sequence 21, Appl
21	41	38.3	39	14	US-10-173-461-65	Sequence 65, Appl
22	41	38.3	71	9	US-09-864-761-43211	Sequence 43211, A
23	41	38.3	80	9	US-09-804-156-38	Sequence 38, Appl
24	41	38.3	80	13	US-10-067-761-38	Sequence 38, Appl
25	41	38.3	80	14	US-10-319-519-38	Sequence 38, Appl
26	41	38.3	153	10	US-09-533-029-96	Sequence 96, Appl
27	41	38.3	153	10	US-09-934-455-40	Sequence 40, Appl
28	41	38.3	153	14	US-10-286-264-108	Sequence 108, App
29	41	38.3	153	14	US-10-278-536-126	Sequence 126, App
30	41	38.3	153	15	US-10-225-068-38	Sequence 38, Appl
31	41	38.3	153	15	US-10-374-780A-2208	Sequence 2208, Ap
32	41	38.3	160	15	US-10-330-051A-44	Sequence 44, Appl
33	41	38.3	254	9	US-09-820-893-94	Sequence 94, Appl
34	41	38.3	264	14	US-10-173-461-2	Sequence 2, Appl
35	41	38.3	315	9	US-09-820-893-59	Sequence 59, Appl
36	41	38.3	324	9	US-09-820-893-96	Sequence 96, Appl
37	41	38.3	327	15	US-10-369-493-3173	Sequence 3173, Ap
38	41	38.3	364	15	US-10-369-493-21445	Sequence 21445, A
39	41	38.3	492	9	US-09-978-295A-7	Sequence 7, Appl
40	41	38.3	492	9	US-09-978-697-7	Sequence 7, Appl
41	41	38.3	492	9	US-09-978-192A-7	Sequence 7, Appl
42	41	38.3	492	9	US-09-999-832A-7	Sequence 7, Appl
43	41	38.3	492	10	US-09-978-189-7	Sequence 7, Appl
44	41	38.3	492	10	US-09-978-608A-7	Sequence 7, Appl
45	41	38.3	492	10	US-09-978-585A-7	Sequence 7, Appl

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1315
US-09-171-432a-39

Query Match      100.0%; Score 107; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLVGKQR 20
Db 1 SHIECRPKYKELRLVGKQR 20

RESULT 2
US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match      100.0%; Score 107; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 4.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLVGKQR 20
Db 164 SHIECRPKYKELRLVGKQR 183

RESULT 3
US-10-272-459-40
; Sequence 40, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-40

Query Match      100.0%; Score 107; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLVGKQR 20
Db 1 SHIECRPKYKELRLVGKQR 20

RESULT 4
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match      100.0%; Score 107; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 1.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLVGKQR 20
Db 792 SHIECRPKYKELRLVGKQR 811

RESULT 5
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 107; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRPKYKELRLVGKQR 20
Db 1 SHIECRPKYKELRLVGKQR 20
```

```
Db      792 SHIECRKPYKELRLEVQKQR 811
|||||
RESULT 6
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 107; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SHIECRKPYKELRLEVQKQR 20
|||||
Db      792 SHIECRKPYKELRLEVQKQR 811
|||||

RESULT 7
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match      100.0%; Score 107; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SHIECRKPYKELRLEVQKQR 20
|||||
Db      792 SHIECRKPYKELRLEVQKQR 811
|||||

RESULT 8
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match      100.0%; Score 107; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SHIECRKPYKELRLEVQKQR 20
|||||
Db      792 SHIECRKPYKELRLEVQKQR 811
|||||

RESULT 9
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match      100.0%; Score 107; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 SHIECRKPYKELRLEVQKQR 20
|||||
Db      792 SHIECRKPYKELRLEVQKQR 811
|||||

RESULT 10
US-09-171-432a-40
; Sequence 40, Application US/0917432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
```

```

US-10-091-007-124
Query Match      42.5%; Score 45.5; DB 14; Length 178;
Best Local Similarity 38.5%; Pred. No. 19;
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

QY      3 IECK-----PYKELELEVGKQ 19
      :||:|||||:|:
Db      103 IKCLKIMAMLRFVFFPKLELRSAGRE 128

RESULT 12
US-09-728-912-13
; Sequence 13, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; TITLE OF INVENTION: and Protein
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 45
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-912-13

Query Match      40.2%; Score 43; DB 9; Length 45;
Best Local Similarity 53.3%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      5 CRKPYKELRLRVGVQ 19
      |||:||||:|
Db      9 CRNPTRELVRGVQ 23

RESULT 13
US-09-728-912-12
; Sequence 12, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; TITLE OF INVENTION: and Protein
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-912-12

Query Match      40.3%; Score 43; DB 9; Length 52;
Best Local Similarity 53.3%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      5 CRKPYKELRLRVGVQ 19
      |||:||||:|
Db      28 CRNPTRELVRGVQ 42

RESULT 14
US-09-728-912-2

```

```

; Sequence 2, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; TITLE OF INVENTION: and Protein
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 413
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-728-912-2

Query Match      40.2%; Score 43; DB 9; Length 413;
Best Local Similarity 53.3%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY      5 CRKPKYKELRLVGVKQ 19
      |||:|:|:|:|
Db      376 CRNPTRELRLVREGVQ 390

RESULT 15
US-09-922-217-1116
; Sequence 1116, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yudiu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1116

Query Match      39.7%; Score 42.5; DB 9; Length 5405;
Best Local Similarity 39.1%; Pred. No. 2.2e+03;
Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

QY      2 HIECRK----PYKELRLVGVKQ 19
      |:|:|:|:|:|:|:|
Db      4824 HVTQEGGAACGPHCECRLEDGVQ 4846

Search completed: March 15, 2004, 13:53:26
Job time : 23.1765 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 15, 2004, 13:49:12 ; Search time 29.1765 Seconds
(without alignments)
216.283 Million cell updates/sec

Title: US-09-171-432A-39
Perfect score: 107
Sequence: 1 SHIECRKPKYKELRLEVGKQR 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_todent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	53	12 Q98VY2	Q98VY2 hepatitis a
2	107	100.0	53	12 Q98VY3	Q98VY3 hepatitis a
3	107	100.0	53	12 Q98VY4	Q98VY4 hepatitis a
4	107	100.0	53	12 Q98VY0	Q98VY0 hepatitis a
5	107	100.0	53	12 Q98VY5	Q98VY5 hepatitis a
6	107	100.0	53	12 Q98VY6	Q98VY6 hepatitis a
7	107	100.0	53	12 Q98VY1	Q98VY1 hepatitis a
8	107	100.0	53	12 Q98VY7	Q98VY7 hepatitis a
9	107	100.0	55	12 Q99VY2	Q99VY2 hepatitis a
10	107	100.0	55	12 Q99VY4	Q99VY4 hepatitis a
11	107	100.0	55	12 Q99VY3	Q99VY3 hepatitis a
12	107	100.0	55	12 Q99VY8	Q99VY8 hepatitis a
13	107	100.0	56	12 Q98VY4	Q98VY4 hepatitis a
14	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
15	107	100.0	56	12 Q98VY2	Q98VY2 hepatitis a
16	107	100.0	56	12 Q99VY1	Q99VY1 hepatitis a

17	107	100.0	56	12 Q99VY8	Q99VY8 hepatitis a
18	107	100.0	56	12 Q98VY2	Q98VY2 hepatitis a
19	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
20	107	100.0	56	12 Q98VY1	Q98VY1 hepatitis a
21	107	100.0	56	12 Q99VY1	Q99VY1 hepatitis a
22	107	100.0	56	12 Q98VY7	Q98VY7 hepatitis a
23	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
24	107	100.0	56	12 Q98VY4	Q98VY4 hepatitis a
25	107	100.0	56	12 Q98VY5	Q98VY5 hepatitis a
26	107	100.0	56	12 Q98VY6	Q98VY6 hepatitis a
27	107	100.0	56	12 Q98VY1	Q98VY1 hepatitis a
28	107	100.0	56	12 Q98VY7	Q98VY7 hepatitis a
29	107	100.0	56	12 Q98VY2	Q98VY2 hepatitis a
30	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
31	107	100.0	56	12 Q98VY4	Q98VY4 hepatitis a
32	107	100.0	56	12 Q98VY5	Q98VY5 hepatitis a
33	107	100.0	56	12 Q98VY6	Q98VY6 hepatitis a
34	107	100.0	56	12 Q98VY1	Q98VY1 hepatitis a
35	107	100.0	56	12 Q98VY7	Q98VY7 hepatitis a
36	107	100.0	56	12 Q98VY2	Q98VY2 hepatitis a
37	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
38	107	100.0	56	12 Q98VY4	Q98VY4 hepatitis a
39	107	100.0	56	12 Q98VY5	Q98VY5 hepatitis a
40	107	100.0	56	12 Q98VY6	Q98VY6 hepatitis a
41	107	100.0	56	12 Q98VY1	Q98VY1 hepatitis a
42	107	100.0	56	12 Q98VY7	Q98VY7 hepatitis a
43	107	100.0	56	12 Q98VY2	Q98VY2 hepatitis a
44	107	100.0	56	12 Q98VY3	Q98VY3 hepatitis a
45	107	100.0	56	12 Q98VY4	Q98VY4 hepatitis a

ALIGNMENTS

RESULT 1

Q98VY2	PRELIMINARY;	PRT;	53 AA.
ID Q98VY2	PRELIMINARY;	PRT;	53 AA.
AC Q98VY2	PRELIMINARY;	PRT;	53 AA.
DT 01-JUN-2001 (Tremblrel. 17, Created)			
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)			
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)			
DE Polyprotein (Fragment).			
OS Hepatitis A virus.			
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC Hepatovirus.			
OX NCBI_TaxID=12092;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN-Uruguay6;			
RA Costa-Mattoli M., Ferre V., Monpocho S., Garcia L., Colina R.,			
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;			
RT "Genetic variability of Hepatitis A Virus in South America reveals			
RT heterogeneity and co-circulation during epidemic outbreaks."			
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR EMBL; AJ309232; CAC37078.1; -			
FT NON_TER	1		
FT NON_TER	53		
SQ SEQUENCE	53 AA;	6271 MW;	0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20

Db 29 SHIECRKPKYKELRLEVGKQR 48

RESULT 2

Q98VY3	PRELIMINARY;	PRT;	53 AA.
ID Q98VY3	PRELIMINARY;	PRT;	53 AA.
AC Q98VY3	PRELIMINARY;	PRT;	53 AA.
DT 01-JUN-2001 (Tremblrel. 17, Created)			

```
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Urul16;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of Hepatitis A Virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309231; CAC37077.1; -.
FT NON_TER 1 1
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPKYKELRLEVGKQR 20
Db 29 SHIECKPKYKELRLEVGKQR 48

RESULT 3
Q98VY4 ID Q98VY4 PRELIMINARY; PRT; 53 AA.
AC Q98VY4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Chile-J;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of Hepatitis A Virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309230; CAC37076.1; -.
FT NON_TER 1 1
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPKYKELRLEVGKQR 20
Db 29 SHIECKPKYKELRLEVGKQR 48

RESULT 4
Q98VY0 ID Q98VY0 PRELIMINARY; PRT; 53 AA.
AC Q98VY0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

```
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Uruguay4;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of Hepatitis A Virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309234; CAC37080.1; -.
FT NON_TER 1 1
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPKYKELRLEVGKQR 20
Db 29 SHIECKPKYKELRLEVGKQR 48

RESULT 5
Q98VY5 ID Q98VY5 PRELIMINARY; PRT; 53 AA.
AC Q98VY5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Urul13;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis A virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ309229; CAC37075.1; -.
FT NON_TER 1 1
FT NON_TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SHIECKPKYKELRLEVGKQR 20
Db 29 SHIECKPKYKELRLEVGKQR 48

RESULT 6
Q98VY6 ID Q98VY6 PRELIMINARY; PRT; 53 AA.
AC Q98VY6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Urul17;
```

RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis A virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ309228; CAC37074.1; -;
FT NON TER 1 1
FT NON TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;
Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKQR 20
DB 29 SHIECRKPKYKELRLVGVKQR 48
RESULT 7
Q98VY1 PRELIMINARY; PRT; 53 AA.
AC Q98VY1; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile1;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis A virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ309233; CAC37079.1; -;
FT NON TER 1 1
FT NON TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;
Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKQR 20
DB 29 SHIECRKPKYKELRLVGVKQR 48
RESULT 8
Q98VY7 PRELIMINARY; PRT; 53 AA.
AC Q98VY7; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Urul;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis A virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ309227; CAC37073.1; -;
FT NON TER 1 1
FT NON TER 53 53
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;
Query Match 100.0%; Score 107; DB 12; Length 53;
Best Local Similarity 100.0%; Pred. No. 3.8e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKQR 20
DB 29 SHIECRKPKYKELRLVGVKQR 48
RESULT 9
Q99T2 PRELIMINARY; PRT; 55 AA.
AC Q99T2; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Urul;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306386; CAC29235.1; -;
DR FIR; PQ0427; PQ0427.
DR FIR; PQ0428; PQ0428.
DR FIR; PQ0430; PQ0430.
DR FIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target_S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON TER 1 1
FT NON TER 55 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;
Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPKYKELRLVGVKQR 20
DB 28 SHIECRKPKYKELRLVGVKQR 47
RESULT 10
Q999T4 PRELIMINARY; PRT; 55 AA.
AC Q999T4; (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-3;
RA Costa-Mattioli M., Ferre V., Monpoeho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic variability of hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks";

```

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306384; CAC29233.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKOR 20
Db 28 SHIECRKPKYKELRLEVGKOR 47

RESULT 11
ID Q999T3 PRELIMINARY; PRT; 55 AA.
AC Q999T3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-9;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
Billaudel S., Vega I., Perez-Barcoff R., Cristina J.;
RT "Genetic variability of Hepatitis C virus reveals heterogeneity and
co-circulation during epidemic outbreaks.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306385; CAC29234.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKOR 20
Db 28 SHIECRKPKYKELRLEVGKOR 47

RESULT 12
ID Q999U8 PRELIMINARY; PRT; 55 AA.
AC Q999U8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg-6;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
Billaudel S., Vega I., Perez-Barcoff R., Cristina J.;
RT "Genetic variability of Hepatitis A Virus in South America reveals
heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306370; CAC29219.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 3.9e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKOR 20
Db 28 SHIECRKPKYKELRLEVGKOR 47

RESULT 13
ID Q8JYP4 PRELIMINARY; PRT; 56 AA.
AC Q8JYP4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Manuveelo;
RA Treambonlers A., Jantaradsamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
outbreak in the southern part of Thailand.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507065; AAM33425.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKOR 20
Db 29 SHIECRKPKYKELRLEVGKOR 48

RESULT 14
ID Q67822 PRELIMINARY; PRT; 56 AA.
AC Q67822;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-70;
RA Karetnyi Y.V., Shulman L.M., Manor J., Leitner L., Shehab S.,
RT Mendelson E.;
RT "Genetic classification of hepatitis A virus strains isolated in
RT Israel, based on their VP1/2A nucleotide sequence";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77247; CAB01040.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56 56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVKGQR 20
DB 29 SHIECRKPKYKELRLEVKGQR 48

RESULT 15
O39872 PRELIMINARY; PRT; 56 AA.
AC O39872;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VDM;
RX MEDLINE=97247817; PubMed=9039340;
RA Taylor N.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RT 1982-1996";
RL J. Med. Virol. 51:273-279(1997).
RL EMBL; U68697; AAB53593.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT CHAIN 1 >56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 4e-10; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVKGQR 20
DB 29 SHIECRKPKYKELRLEVKGQR 48

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:23:11 ; Search time 47 Seconds
(without alignments)
120.233 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVQKRLKYAEE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_29Jan04.*

1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	20	2	AAW42923 Immunogen
2	102	100.0	21	4	AAW69440 Synthetic
3	102	100.0	366	1	AAP50230 Sequence
4	102	100.0	854	1	AAP50287 Hepatitis
5	102	100.0	993	1	AAP50116 Sequence
6	102	100.0	1077	2	AAP50231 Sequence
7	102	100.0	1092	2	AAW95559 A partial
8	102	100.0	1091	2	AAW95559 A partial
9	102	100.0	2227	2	AAW34074 Hepatitis
10	102	100.0	2227	2	AAW34074 Hepatitis
11	102	100.0	2227	3	AAW18609 Amino aci
12	102	100.0	2227	3	AAW18609 Amino aci
13	102	100.0	2227	3	AAW18609 Amino aci
14	102	100.0	2227	3	AAW18609 Amino aci
15	102	100.0	2227	5	AAW18609 Amino aci
16	102	100.0	2227	5	AAW18609 Amino aci
17	102	100.0	2227	5	AAW18609 Amino aci
18	102	100.0	2227	6	AAW18609 Amino aci
19	102	100.0	2227	6	AAW18609 Amino aci
20	102	100.0	2227	6	AAW18609 Amino aci
21	102	100.0	2227	7	AAW18609 Amino aci
22	99	97.1	2227	7	AAW18609 Amino aci
23	94	92.2	839	2	AAW15629 Capsid re
24	67	65.7	20	2	AAW42922 Immunogen
25	67	65.7	21	4	AAW69439 Synthetic

99 41 40.2 859 6 ABU38058 Protein e
100 40.5 39.7 360 4 ABG00270 Novel hum

ALIGNMENTS

```

RESULT 1
AAW42923
ID AAW42923 standard; peptide; 20 AA.
XX
XX
AC AAW42923;
XX
XX
DT 28-APR-1998 (first entry)
XX
XX
DE Immunogenic Hepatitis A virus peptide YK-1316.
XX
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX
XX
OS Synthetic.
OS Hepatitis A virus.
XX
XX
FN WO9740147-A1.
XX
XX
PD 30-OCT-1997.
XX
XX
PF 18-APR-1997; 97WO-US006891.
XX
XX
PR 19-APR-1996; 96US-0015644P.
XX
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Fields HA, Khudyakov YE;
XX
XX
DR WPI; 1997-535831/49.
XX
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
XX
PS Claim 18; Page 112; 140pp; English.
XX
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 799-818, and has a reactivity of 41.7% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal
XX
XX
SQ Sequence 20 AA;

Query Match 100.0%; Score 102; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 1 PYKELRLEVGKQRLKYAQEE 20

RESULT 2
AAB69440
ID AAB69440 standard; peptide; 21 AA.
XX
XX
AC AAB69440;
XX
XX
DT 20-APR-2001 (first entry)
XX

```

```

DE Synthetic HAV P2A peptide, SEQ ID NO: 40.
XX
XX
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
XX
OS Hepatitis A virus.
OS Synthetic.
XX
XX
FN WO200105824-A2.
XX
XX
PD 25-JAN-2001.
XX
XX
PF 14-JUL-2000; 2000WO-US019267.
XX
XX
PR 15-JUL-1999; 99US-0144412P.
XX
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX
PI Fields HA, Khudyakov YE;
XX
XX
DR WPI; 2001-112681/12.
XX
XX
PT Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
XX
XX
PS Claim 13; Page 93; 130pp; English.
XX
XX
CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity
XX
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 102; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 1 PYKELRLEVGKQRLKYAQEE 20

RESULT 3
AAP50230
ID AAP50230 standard; protein; 366 AA.
XX
XX
AC AAP50230;
XX
XX
DT 28-NOV-1991 (first entry)
XX
XX
DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX
XX
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
XX
OS Hepatitis A virus.
XX
XX
FN EPI38704-A.
XX
XX
PD 24-APR-1985.

```

XX PF 09-OCT-1984; 84EP-00402025.
XX PR 14-OCT-1983; 83US-00541836.
XX PR 02-MAR-1984; 84US-00585942.
XX PA (MERI) MERCK & CO INC.
XX PI Hughes JV, Scolnick EM, Tomassini JE;
XX DR WPI; 1985-100818/17.
XX DR N-PSDB; AAN50274.
XX PT New hepatitis A virus surface protein - useful for binding to
XX PT neutralising antibodies to the virus.
XX PS Claim 21; Page 46-48; 49pp; English.
XX CC VP1 is isolated by solubilisation of the intact virus in an aq. anionic
XX CC surfactant and a reducing agent. The viral proteins are sepd. and the
XX CC protein of molecular wt. 33000 daltons is sepd
XX SQ Sequence 366 AA;
Query Match 100.0%; Score 102; DB 1; Length 366;
Best Local Similarity 100.0%; Pred. No. 3.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PYKELELVGKQRLKYAOEE 20
Db 308 PYKELELVGKQRLKYAOEE 327
RESULT 4
AAP50287
ID AAP50287 standard; protein; 854 AA.
XX AC AAP50287;
XX DT 25-MAR-2003 (revised)
XX DT 30-NOV-1991 (first entry)
XX DE Hepatitis A virus (HAV) peptide corresponding to the capsid protein
XX DE region of poliovirus RNA.
XX KW Hepatitis A virus assay; antigen; antibody.
XX OS Hepatitis A virus.
XX PN WO8501517-A.
XX PD 11-APR-1985.
XX PF 27-SEP-1984; 84WO-US001552.
XX PR 30-SEP-1983; 83US-00537911.
XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
XX PI Racaniello VR;
XX DR WPI; 1985-098846/16.
XX DR N-PSDB; AAN50330.
XX PT New hepatitis A virus CDNA - useful in assays for the virus and for
XX PT prodn. of the viral antigen and antibodies to it.
XX PS Example; Fig 7; 60pp; English.
XX CC The inventors claim HAV cDNA and a method for producing it, whereby large
XX CC amts. can be obtd. economically. The cDNA is useful in the assay for
XX CC detection of HAV quickly and easily and with high sensitivity and

CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
CC 2003 to correct PA field.)
XX SQ Sequence 854 AA;
Query Match 100.0%; Score 102; DB 1; Length 854;
Best Local Similarity 100.0%; Pred. No. 9.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PYKELELVGKQRLKYAOEE 20
Db 799 PYKELELVGKQRLKYAOEE 818
RESULT 5
AAP50116
ID AAP50116 standard; protein; 993 AA.
XX AC AAP50116;
XX DT 24-OCT-2003 (revised)
XX DT 25-MAR-2003 (revised)
XX DT 30-SEP-1991 (first entry)
XX DE Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
XX DE and VP-4.
XX KW Antigenic protein; immunogen; vaccine.
XX OS Hepatitis A virus; (strain CR326).
XX PN EP154587-A.
XX PD 11-SEP-1985.
XX PF 27-FEB-1985; 85EP-00400369.
XX PR 02-MAR-1984; 84US-00585818.
XX PA (MERI) MERCK & CO INC.
XX PI Lineneyer DL, Menke JG, Rueben RG, Mitra SM;
XX DR WPI; 1985-224964/37.
XX DR N-PSDB; AAN50139.
XX PT New nucleotide sequences coding for hepatitis A virus antigens - useful
XX PT for eliciting normal immune response and in vaccines for protecting
XX PT against the virus.
XX PS Example; Page 11-17; 32pp; English.
XX CC Within the sequence in AAN50139 is encoded the information necessary to
XX CC make the antigenic proteins of HAV. The sequences encoding for the
XX CC structural proteins begin at base 403. The key sub-unit sequences within
XX CC VP-1, designated Sequences I, II, III, IV, and V, start, respectively at
XX CC 1882, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
XX CC valuable as encoding antigenic proteins are the sequences from base 1749
XX CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
XX CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
XX CC vector for producing antigen protein. Sequences II-V are claimed. X in
XX CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
XX CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 993 AA;
Query Match 100.0%; Score 102; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PYKELELVGKQRLKYAOEE 20
Db 799 PYKELELVGKQRLKYAOEE 818

Db 935 PYKELRLEVGKQRLKYAQEE 954

RESULT 6
AAP50231
ID AAP50231 standard; protein; 993 AA.
XX AC AAP50231;
XX 28-NOV-1991 (first entry)
DT DT
XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX Hepatitis A virus.
OS
XX Key Location/Qualifiers
FH 628..993
FT Protein
FT /note= "clained; X denotes translated stop codons and
FT unspecified triplets"
XX EPI38704-A.
XX 24-APR-1985.
XX 09-OCT-1984; 84EP-00402025.
XX 14-OCT-1983; 83US-00541836.
XX 02-MAR-1984; 84US-00585942.
XX (MERI) MERCK & CO INC.
XX Hughes JV, Scolnick EM, Tomassini JE;
XX WPI; 1985-100818/17.
XX N-PSDB; AAN50274.
XX New hepatitis A virus surface protein - useful for binding to
PT neutralising antibodies to the virus.
XX Disclosure; Page 17-23; 49pp; English.
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 daltons is sepd
XX Sequence 993 AA;
XX Query Match 100.0%; Score 102; DB 1; Length 993;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 935 PYKELRLEVGKQRLKYAQEE 954

RESULT 7
AAW95559
ID AAW95559 standard; protein; 1077 AA.
XX AC AAW95559;
XX 28-APR-1999 (first entry)
DT A partial hepatitis A virus (HAV) protein.
XX Hepatitis A virus protein; HAV; P2 region;
KW cell-culture-adapted HAV strain; infection; accelerated growth.
XX

OS Hepatitis A virus.
XX US5849562-A.
XX 15-DEC-1998.
XX 06-JUN-1995; 95US-00468926.
XX 30-SEP-1983; 83US-00537911.
PR 27-SEP-1984; 84US-00654942.
PR 06-OCT-1988; 88US-00256135.
PR 06-NOV-1991; 91US-00788262.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA Emerson SU, Purcell RH;
XX WPI; 1999-094412/08.
DR N-PSDB; AAX01006.
XX Chimeric hepatitis A virus strains - with P2 region from cell-culture-
PT adapted strain in wild-type genome.
XX Disclosure; Fig 7A-L; 36pp; English.
XX The present sequence represents a partial hepatitis A virus (HAV)
CC protein. The specification describes a DNA construct consisting of a wild
CC -type HAV genome in which the P2 region is replaced by the P2 region from
CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
CC sufficient for establishment of infection and accelerated growth in cell
CC culture
XX Sequence 1077 AA;
SQ Query Match 100.0%; Score 102; DB 2; Length 1077;
XX Best Local Similarity 100.0%; Pred. No. 1.2e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 1022 PYKELRLEVGKQRLKYAQEE 1041

RESULT 8
AAR32426
ID AAR32426 standard; protein; 1091 AA.
XX AC AAR32426;
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 10-JUN-1993 (first entry)
XX Translated from 5' region of Hepatitis A Virus genomic clone.
DE HAV HM-175; chronic liver disease; picornavirus.
XX Hepatitis A virus.
OS
XX Key Location/Qualifiers
FH 1..711
FT /note= "X's correspond to nonsense codons, i.e. this
FT region is not an ORF"
FT 238..1091
FT /label= ORF
FT /note= "second putative initiation codon at position 240"
XX USN7788262-N.
XX 15-DEC-1992.
PD

PF 06-NOV-1991; 91US-00788262.
 XX 30-SEP-1983; 88US-00536911.
 PR 27-SEP-1984; 84US-00654942.
 PR 06-OCT-1988; 88US-00256135.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 XX Racaniello VR, Baroudy BM, Emerson SU;
 PI WPI; 1993-067429/08.
 DR N-PSDB; AAQ36934.
 XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
 PT antigen and antibodies.
 PT Disclosure; Fig 7; 65pp; English.
 PS HAV virion RNA was extracted from the livers of marmosets which had been
 XX inoculated with HAV (the HAV had previously been passaged twice in
 CC marmosets). The RNA was used to prepare ds cDNA clones by standard
 CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected
 CC African Green Monkey Kidney cells were selected for further analysis. A
 CC 7.4kb restriction map (about 99% of the HAV genome) was constructed from
 CC 5 overlapping inserts. The sequence of the first 3.3kb (approx.) from the
 CC 5'-terminus was determined. An amino acid sequence was deduced from the
 CC entire clone and an open reading frame was identified starting at
 CC position 238. A comparison of the predicted HAV amino acid sequences with
 CC the known capsid protein sequences of other picornaviruses (poliovirus,
 CC foot and mouth disease virus and encephalomyelitis virus) revealed areas
 CC of local homology. (Note: Revised entry submitted to correct the patent
 CC number format of US Government-owned NTIS applications to prevent clashes
 CC with ongoing US granted patent numbers. For further information please
 CC visit the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html.)
 CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX SQ Sequence 1091 AA;
 Query Match 100.0%; Score 102; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 1036 PYKELRLEVGKQRLKYAQEE 1055
 RESULT 9
 ID AAR05697
 XX AAR05697 standard; protein; 2227 AA.
 XX AAR05697;
 XX 24-OCT-2003 (revised)
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 15-AUG-1990 (first entry)
 XX Attenuated hepatitis A virus.
 DE Hepatitis A virus; vaccine; attenuated.
 KW Hepatitis A virus; strain HM-175.
 OS Hepatitis A virus; protein; 2227 AA.
 XX Key Location/Qualifiers
 FH 1..23
 FT Region /label= VP4 = 1A
 FT Region 24..245
 FT /label= VP2 = 1B
 FT 246..491
 FT /label= VP3 = 1C

Region 492..791
 FT /label= VP1 = 1D
 FT 792..980
 FT /label= 2A
 FT 981..1087
 FT /label= 2B
 FT 1088..1422
 FT /label= 2C
 FT 1423..1496
 FT /label= 3A
 FT 1497..1519
 FT /label= 3B = VPg
 FT 1520..1738
 FT /label= 3C
 FT 1739..2227
 FT /label= 3D
 XX US4894228-A.
 PN 16-JAN-1990.
 XX 12-JUL-1988; 88US-00217824.
 XX 19-SEP-1984; 84US-00652067.
 PR 09-SEP-1986; 86US-00905146.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
 PI Raemer RJ, Gust ID;
 XX WPI; 1990-075557/10.
 DR N-PSDB; AAQ03512.
 XX Vaccine against hepatitis A virus infection - comprises novel attenuated
 PT hepatitis A virus strain.
 XX Claim 1; Fig 1; 18pp; English.
 PS The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome. It
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
 CC standardise OS field)
 XX SQ Sequence 2227 AA;
 Query Match 100.0%; Score 102; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 799 PYKELRLEVGKQRLKYAQEE 818
 RESULT 10
 ID AAW34074
 XX AAW34074 standard; protein; 2227 AA.
 XX AAW34074;
 XX 17-OCT-2003 (revised)
 DT 27-APR-1998 (first entry)
 XX Hepatitis A virus HM-175 protein sequence.
 DE HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
 KW vaccine.

XX Hepatitis A virus; HM-175.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Protein 1. .23
 FT Protein /label= VP4
 FT Protein 24. .245
 FT Protein /label= VP2
 FT Protein 246. .491
 FT Protein /label= VP3
 FT Protein 492. .791
 FT Protein /label= VP1
 FT Protein 792. .980
 FT Protein /label= 2A
 FT Protein 981. .1087
 FT Protein /label= 2B
 FT Protein 1088. .1422
 FT Protein /label= 2C
 FT Protein 1423. .1496
 FT Protein /label= 3A
 FT Protein 1497. .1519
 FT Protein /label= 3B
 FT Protein 1520. .1738
 FT Protein /label= 3C
 FT Protein 1739. .2227
 FT Protein /label= 3D
 FT
 PN WO9740166-A2.
 XX
 XX 30-OCT-1997.
 PD
 XX
 XX 18-APR-1997; 97WO-US006506.
 PF
 XX
 XX 19-APR-1996; 96US-0015642P.
 PR
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX
 XX Raychaudhuri G, Emerson SU, Purcell RH;
 PI
 XX WPI; 1997-535850/49.
 DR
 DR N-PSDB; AAT93023.
 XX
 XX Human attenuated HAV genome containing simian HAV 2C gene - useful as
 PT vaccines against HAV infection.
 PT
 XX
 XX Disclosure; Fig 13A-D; 66pp; English.
 PS
 XX
 XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
 CC 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
 CC by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
 CC construct (I) comprises a genome of HAV, where the genome is a human
 CC attenuated HAV genome in which a region of the 2C gene has been replaced
 CC by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
 CC (see AAT93024). The region of the 2C gene from AGM-27 contained in the
 CC construct preferably encodes amino acids 120-328 of the 2C protein, amino
 CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
 CC of (I); (2) a cell transfected with (I) or the RNA transcript of (1); (3)
 CC a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
 CC cell containing the HAV of (3). (I) or its RNA transcript, can be used as
 CC a vaccine for preventing HAV in a mammal. (I) or the RNA transcript can
 CC also be used to stimulate the production of protective antibodies in the
 CC mammal. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 102; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 11
 AAB18609
 ID AAB18609 standard; protein; 2227 AA.
 AC
 XX AAB18609;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of live attenuated Hepatitis A virus 4380.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 KW HAV 4380.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.
 XX
 PD 05-SEP-2000.
 PF 07-JUN-1995; 95US-00475886.
 PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 PI
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75478.
 DR
 XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.
 XX
 XX Disclosure; Col 93-104; 72pp; English.
 PS
 XX The present sequence is derived from a live attenuated hepatitis A virus
 CC (HAV) of the invention, designated HAV 4380. The sequence is produced by
 CC modifying wild type HAV strain HM-174. The HAV of the invention are
 CC adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
 CC is able to propagate in MRC-5 cells and retain appropriate attenuation.
 CC It is useful as a live vaccine for prophylaxis of hepatitis A in humans
 CC and other primates
 CC
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 102; DB 3; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 799 PYKELRLEVGKQRLKYAQEE 818
 RESULT 12
 AAB18607
 ID AAB18607 standard; protein; 2227 AA.
 AC
 XX AAB18607;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
 XX
 OS Hepatitis A virus.
 XX
 PN US6113912-A.

XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX PR 18-SEP-1992; 92US-00947338.
 XX PR 17-SEP-1993; 93WO-US008610.
 XX PR 17-APR-1995; 95US-00397232.
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX DR WPI; 2000-586464/55.
 XX DR N-PSDB; AAA75476.
 XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
 XX PT line useful as vaccine for protecting humans against hepatitis A virus
 XX PT infection, has modified genome compared to wild type.
 XX PS Disclosure; Fig 6A-K; 72pp; English.
 XX CC The present sequence is derived from a wild type hepatitis A virus (HAV)
 XX CC strain HM-174. The sequence is modified to produce HAV which are adapted
 XX CC to growth in the human fibroblast-like cell line MRC-5. The HAV is able
 XX CC to propagate in MRC-5 cells and retain appropriate attenuation. It is
 XX CC useful as a live vaccine for prophylaxis of hepatitis A in humans and
 XX CC other primates
 XX CC
 XX CC Sequence 2227 AA;
 XX CC
 XX CC Query Match 100.0%; Score 102; DB 3; Length 2227;
 XX CC Best Local Similarity 100.0%; Pred. No. 2,7e-07;
 XX CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX CC
 XX QY 1 PYKELRLEVGKQRLKYAQEE 20
 XX DB |||||
 XX DB 799 PYKELRLEVGKQRLKYAQEE 818
 XX
 XX RESULT 13
 XX AAB18608
 XX ID AAB18608 standard; protein; 2227 AA.
 XX AC AAB18608;
 XX XX
 XX DT 15-JAN-2001 (first entry)
 XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
 XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 XX KW P-35 virus.
 XX OS Hepatitis A virus.
 XX PN US6113912-A.
 XX PD 05-SEP-2000.
 XX PF 07-JUN-1995; 95US-00475886.
 XX PR 18-SEP-1992; 92US-00947338.
 XX PR 17-SEP-1993; 93WO-US008610.
 XX PR 17-APR-1995; 95US-00397232.
 XX XX
 XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX DR WPI; 2000-586464/55.
 XX DR N-PSDB; AAA75477.
 XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.
 XX Disclosure; Col 67-78; 72pp; English.
 XX CC The present sequence is derived from passage 35 of a wild type hepatitis
 XX CC A virus (HAV) strain HM-174. The resulting virus is designated P-35
 XX CC virus. The sequence is modified to produce HAV which are adapted to
 XX CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
 XX CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
 XX CC as a live vaccine for prophylaxis of hepatitis A in humans and other
 XX CC primates
 XX CC
 XX CC Sequence 2227 AA;
 XX CC
 XX CC Query Match 100.0%; Score 102; DB 3; Length 2227;
 XX CC Best Local Similarity 100.0%; Pred. No. 2,7e-07;
 XX CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX CC
 XX QY 1 PYKELRLEVGKQRLKYAQEE 20
 XX DB |||||
 XX DB 799 PYKELRLEVGKQRLKYAQEE 818
 XX
 XX RESULT 14
 XX AAE19899
 XX ID AAE19899 standard; protein; 2227 AA.
 XX AC AAE19899;
 XX XX
 XX DT 18-JUN-2002 (first entry)
 XX DE Hepatitis A virus (HAV) protein.
 XX KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
 XX KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
 XX OS Hepatitis A virus.
 XX PN WO200213855-A2.
 XX PD 21-FEB-2002.
 XX PF 15-AUG-2001; 2001WO-IB001808.
 XX PR 17-AUG-2000; 2000US-0225767P.
 XX PR 29-AUG-2000; 2000US-0229175P.
 XX PR 03-NOV-2000; 2000US-00705547.
 XX XX
 XX FA (TRIP-) TRIPEP AB.
 XX PI Sallberg M, Hultgren C;
 XX DR WPI; 2002-241837/29.
 XX DR N-PSDB; AAD31766.
 XX PT Vaccine compositions for treating and preventing disease, preferably
 XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
 XX PT epitope present in hepatitis C virus.
 XX PS Claim 11; Page 82-87; 120pp; English.
 XX CC The invention relates to a composition comprising ribavirin and an
 XX CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 XX CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 XX CC sequence. The composition is useful for enhancing an immune response to a
 XX CC hepatitis C antigen in humans, domestic, sport or pet species and as
 XX CC vaccines for treating and preventing HCV infections. The composition is
 XX CC also useful for treating viral, bacterial, fungal diseases and cancer.
 XX CC The present sequence is hepatitis A virus (HAV) protein
 XX CC
 XX CC Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVQKQRLKYAQEE 20
Db 799 PYKELRLEVQKQRLKYAQEE 818

RESULT 15
ABG31729
ID ABG31729 standard; protein; 2227 AA.
XX AC ABG31729;
XX 29-AUG-2003 (revised)
DT 29-NOV-2002 (first entry)
XX 29-AUG-2003 (revised)
DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
KW HAV 4380.
XX Hepatitis A virus; strain HM-175.
OS Hepatitis A virus; strain HM-175.
XX US6423318-B1.
PN 23-JUL-2002.
XX 31-AUG-2000; 2000US-00653499.
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
XX N-PSDB; ABS52789.
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX Disclosure; Col 93-104; 71pp; English.
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents an attenuated
CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
CC standardise OS field)
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVQKQRLKYAQEE 20
Db 799 PYKELRLEVQKQRLKYAQEE 818

RESULT 16
ABG31727
ID ABG31727 standard; protein; 2227 AA.
XX AC ABG31727;
XX 29-AUG-2003 (revised)
DT 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)
XX Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX Hepatitis A virus; strain HM-175.
XX US6423318-B1.
XX 23-JUL-2002.
XX 31-AUG-2000; 2000US-00653499.
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2002-680946/73.
XX N-PSDB; ABS52789.
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX Disclosure; Fig 6; 71pp; English.
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
CC field)
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 102; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVQKQRLKYAQEE 20
Db 799 PYKELRLEVQKQRLKYAQEE 818

RESULT 17
ABG31728
ID ABG31728 standard; protein; 2227 AA.
XX AC ABG31728;
XX 29-NOV-2002 (first entry)
XX Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
KW virucide; mutant; pHAV/7; mutein.
XX Hepatitis A virus; strain HM-175.
OS Synthetic.
XX Key Location/Qualifiers
PH Misc-difference 764
FT /note= "Wild-type Glu substituted by Val"
FT Misc-difference 821
FT /note= "Wild-type Asn substituted by Ser"
FT Misc-difference 963
FT /label= Wild-type Lys substituted by Arg
FT Misc-difference 1052

FT FT /note= "Wild-type Ala substituted by Val"
Misc-difference 1062
FT /note= "Wild-type Gly substituted by Ala"
Misc-difference 1118
FT /note= "Wild-type Lys substituted by Met"
Misc-difference 1151
FT /note= "Wild-type Glu substituted by Lys"
Misc-difference 1163
FT /note= "Wild-type Phe substituted by Ser"
Misc-difference 1277
FT /note= "Wild-type Val substituted by Ile"
Misc-difference 1500
FT /note= "Wild-type His substituted by Tyr"
Misc-difference 1805
FT /note= "Wild-type Asp substituted by Asn"
Misc-difference 1930
FT /note= "Wild-type Ser substituted by Thr"
XX
PN US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-00653499.
XX
XX 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
PI
XX
XX WPI; 2002-680946/73.
DR N-PSDB; ABS52788.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
PT
XX Example 3; Col 67-78; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC mutant strain HM-175/7 (pHAV/7) polypeptide
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 102; DB 5; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 799 PYKELRLEVGKQRLKYAQEE 818
XX
XX RESULT 18
ABU08640
ID ABU08640 standard; protein; 2227 AA.
XX
XX ABU08640;
XX
XX 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
XX Attenuated (pass35) hepatitis A virus strain HM-175.
DE
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine, MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
XX Hepatitis A virus; strain HM-175.
OS

XX US2002176869-A1.
PN
XX 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
PF
XX 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
PI
XX
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93474.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
XX Example 3; Fig 6; 70pp; English.
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 102; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 799 PYKELRLEVGKQRLKYAQEE 818
XX
XX RESULT 19
ABU08641
ID ABU08641 standard; protein; 2227 AA.
XX
XX ABU08641;
XX
XX 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
XX Attenuated hepatitis A virus (4380) strain HM-175.
DE
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine, MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
XX Hepatitis A virus; strain HM-175.
XX
XX US2002176869-A1.
XX
XX 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
PF
XX 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR

XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HOND T E.
XX
XX Funkhouser AW, Emerson SU, Furcell RH, D'hondt E;
PI
XX
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93473.
DR

XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis C virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis A
CC virus protein
CC
CC
CC Sequence 2227 AA;
SO

```

Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PYKELRLEVGKQRLKYAQEE 20
      |||||
Db      799 PYKELRLEVGKQRLKYAQEE 818

RESULT 22
AAP60066
ID AAP60066 standard; protein; 2227 AA.
XX
AC AAP60066;
XX
DT 25-MAR-2003 (revised)
DT 26-JUN-1991 (first entry)
XX
DE Sequence of viral L434 polypeptide encoded by the complete nucleotide
DE sequence of the HAV genome.
XX
KW Diagnosis; vaccine; passive immunotherapy.
XX
OS Hepatitis A virus.
XX
FH Key Location/Qualifiers
FH Region 1..245
FT FT /label= P1.1A
FT FT 246..491
FT FT /label= 1B
FT FT 492..836
FT FT /label= 1C
FT FT 837..980
FT FT /label= P2.2A
FT FT 981..1076
FT FT /label= 2B
FT FT 1077..1422
FT FT /label= 2C
FT FT 1423..1484
FT FT /label= P3.3A
FT FT 1485..1507
FT FT /label= 3B
FT FT 1508..1678
FT FT /label= 3C
FT FT 1679..2227
FT FT /label= 3D
XX
PN EP199480-A.
XX
PD 29-OCT-1986.
XX
PF 03-APR-1986; 86EP-00302465.
XX
PR 03-APR-1985; 85US-00719329.
XX
      (CHIR ) CHIRON CORP.
XX
PI Dina D, Potter SJ, Vannest GA, Caput D;
XX
WP1; 1986-286213/44.
DR N-PSDB; AAP60080.
DR
XX
PT Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.
PT of vaccines and diagnostic probes.
XX
XX Claim 5; Fig 1; 18pp; English.
XX
XX AAP60080 and oligonucleotide fragments are useful in detection of
XX hepatitis A virus; transformed hosts may be used for expression of
XX polypeptides and fragments useful in vaccines without risk of infection
XX by the virus or in prodn. of particles which are capable of inducing
XX immunocompetent B cells for passive immunotherapy. Pref. epitope is
XX derived from AAs 445-657 or 792-948 of the HAV polypeptide sequence
XX (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)
XX

```


CC cyto-HAV. The other two cleavage sites are the same. Two residues have
CC been identified as part of the immuno- dominant region (see feature
CC table) and are different to those in the same position in human HAV. The
CC protein and peptides derived from it can be used in the prepn. of
CC vaccines for the prevention of HAV infection. See also AAR15056. (Note:
CC Revised entry submitted to correct the patent number format of US
CC Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-
CC MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 839 AA;

Query Match 92.2%; Score 94; DB 2; Length 839;
Best Local Similarity 90.0%; Pred. No. 1.9e-06;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKQRLKYAQEE 20
DB 798 PYKELRLVGVKQRFKYAREE 817
|||||

RESULT 24
AAW42922
ID AAW42922 standard; peptide; 20 AA.
XX
AC AAW42922;
XX

DT 28-APR-1998 (first entry)
XX

XX Immunogenic Hepatitis A virus peptide YK-1315.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.

OS Synthetic.
OS Hepatitis A virus.

PN WO9740147-A1.
XX

PD 30-OCT-1997.
XX

PF 18-APR-1997; 97WO-US006891.
XX

PR 19-APR-1996; 96US-0015644P.
XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

XX Fields HA, Khudyakov YE;
XX

XX WPI; 1997-535831/49.
XX

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.

XX Claim 18; Page 112; 140pp; English.
XX

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 792-811, and has a reactivity of 54.2% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal
XX

SQ Sequence 20 AA;

Query Match 65.7%; Score 67; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.00086;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKQRL 13
|||||

DB 8 PYKELRLVGVKQRL 20
|||||

RESULT 25
AAB69439
ID AAB69439 standard; peptide; 21 AA.
XX
AC AAB69439;
XX

DT 20-APR-2001 (first entry)
XX

XX Synthetic HAV P2A peptide, SEQ ID NO: 39.

XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.

XX Hepatitis A virus.
OS Synthetic.

XX WO200105824-A2.
PN

XX 25-JAN-2001.
PD

PF 14-JUL-2000; 2000WO-US019267.
XX

PR 15-JUL-1999; 99US-0144412P.
XX

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX

XX Fields HA, Khudyakov YE;
XX

XX WPI; 2001-112681/12.
DR

XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.

XX Claim 13; Page 93; 130pp; English.
XX

XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity
XX

SQ Sequence 21 AA;

Query Match 65.7%; Score 67; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 0.00091;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKQRL 13
|||||

DB 8 PYKELRLVGVKQRL 20
|||||

RESULT 26
ADA33461

ADA33461 standard; protein; 249 AA.
 ADA33461;
 20-NOV-2003 (first entry)
 Acinetobacter baumannii protein #622.
 Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;
 plant biocontrol agent.
 Acinetobacter baumannii.
 US6562958-B1.
 13-MAY-2003.
 04-JUN-1999; 95US-00328352.
 09-JUN-1998; 98US-0088701P.
 (GENO-) GENOME THERAPEUTICS CORP.
 Breton G, Bush D;
 WPI; 2003-576092/54.
 N-PSDB; ADA29335.
 New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
 for diagnosing a bacterial disease, as components of antibacterial
 vaccines, as targets for antibacterial drugs, or as biocontrol agents for
 plants.
 Example; SEQ ID NO 4748; 328pp; English.
 The invention relates to isolated Acinetobacter baumannii nucleic acids.
 The A. baumannii nucleic acids and polypeptides are useful as reagents
 for diagnosing a bacterial disease, as components of antibacterial
 vaccines, as targets for antibacterial drugs, to detect the presence of
 A. baumannii and other Acinetobacter species in a sample, in screening
 compounds for the ability to interfere with the A. baumannii life cycle
 or to inhibit A. baumannii infection, and as biocontrol agents for
 plants. The present sequence represents the amino acid sequence of an A.
 baumannii protein.
 Sequence 249 AA;
 Query Match 46.1%; Score 47; DB 6; Length 249;
 Best Local Similarity 56.2%; Pred. No. 22;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 1 PYKLELVGKQRLKY 16
 DB 5 PMKSLWLDIGNTRUKY 20
 RESULT 27
 ABP31097
 ID ABP31097 standard; protein; 65 AA.
 AC ABP31097;
 08-JUL-2002 (first entry)
 Human ORF70 protein, SEQ ID NO:140.
 Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 disease monitoring; cytokine; cell proliferation; cell differentiation;
 immune modulation; haematopoiesis regulation; tissue growth;
 angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
 thrombolytic; tumour inhibition; bodily characteristic; fertility;
 behaviour; cancer; proliferative disorder; neurological disorder;
 cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 hypothyroidism; cholesterol ester storage disease; infection; vulvular;
 vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 dermatological; analgesic; virucide; antibacterial; fungicide.
 Homo sapiens.
 WO200190366-A2.
 29-NOV-2001.
 24-MAY-2001; 2001WO-US017076.
 24-MAY-2000; 2000US-0206690P.
 (CURA-) CURAGEN CORP.
 Leach MD, Shinkets RA;
 WPI; 2002-106200/14.
 N-PSDB; ABN75123.
 Novel human polypeptides and polynucleotides useful for diagnosing,
 preventing and treating cardiovascular disease, neurodegenerative,
 hyperproliferative disorders and disorders related to organ
 transplantation.
 Claim 10; Page 305; 2508pp; English.
 Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 ABN79587 represent cDNAs encoding them. The invention also encompasses
 polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 referred to as ORFX) proteins, polynucleotides at least 85% identical to
 the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 polynucleotides, the recombinant production of ORFX proteins, antibodies
 specific for ORFX proteins, methods of detecting ORFX polynucleotides or
 polypeptides, methods of screening for modulators of ORFX expression or
 activity, and methods of screening individuals for a predisposition to an
 ORFX-associated disorder. The ORFX proteins of the invention have a wide
 range of biological activities, such as cytokine, cell proliferation,
 cell differentiation, immune modulation, haematopoiesis regulation,
 tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 chemokinetic activity, haemostatic activity, thrombolytic activity,
 receptor/ligand, antiinflammatory activity, tumour inhibition activity,
 and antiinfective activity, and may also be involved in the determination
 of bodily characteristics, fertility and behaviour. ORFX proteins,
 nucleic acids and antibodies may be used in the treatment of cancers,
 other proliferative disorders such as psoriasis and benign tumours,
 neurological disorders such as epilepsy and Alzheimer's disease,
 cardiovascular diseases, immune system disorders, disorders related to
 organ transplantation, disorders of tissue growth and regeneration,
 diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 storage disease, and infectious diseases caused by viral, bacterial,
 fungal and other pathogens. ORFX nucleic acids may also be used as a
 source of primers and probes, in the detection of ORFX genomic sequences
 or transcripts, in the identification and cloning of homologous
 sequences, in genetic diagnosis, and in forensic biology. The ORFX
 nucleic acids may additionally be used to produce transgenic animals
 which may be useful for studying the function and/or activity of ORFX
 protein, and in drug screening. The ORFX proteins may also be used as
 immunogens to generate specific antibodies, which are useful in the
 diagnosis, treatment and monitoring of ORFX-associated diseases
 XX
 SQ Sequence 65 AA;
 Query Match 45.6%; Score 46.5; DB 5; Length 65;
 Best Local Similarity 55.6%; Pred. No. 6.4;
 Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
 QY 1 PYKLELVGKQRLKYAQ 18

Wed May 12 03:35:21 / 2004

Db Db

34 PYKGLR-EISRNLRYAQ 50

RESULT 28
ABP65745
ID ABP65745 standard; protein; 375 AA.
XX AC
XX ABP65745;
DT 19-NOV-2002 (first entry)
DE Bifidobacterium longum NCC2705 ORF amino acid sequence SEQ ID NO:499.
XX DE
XX Bifidobacterium longum NCC2705; Bifidobacterium; bacterial;
KW antidiarrheic; antibacterial; inhibitor of Salmonella; detection;
KW identification; lactic acid bacterium; diarrhoea; pathogenic bacteria;
KW rotavirus; food composition; pharmaceutical composition.
XX KW
XX Bifidobacterium longum.
OS OS
EP1227152-A1.
XX EP
XX 31-JUL-2002.
XX PD
XX 30-JAN-2001; 2001EP-00102050.
XX PF
XX 30-JAN-2001; 2001EP-00102050.
XX PR
XX (NEST) SOC PROD NESTLE SA.
PA PA
WPI; 2002-668397/72.
DR DR
XX The present invention describes a polynucleotide (I) comprising a
sequence of a Bifidobacterium genome selected from the nucleotide
sequences given in ABQ81842 and ABQ81843, or a sequence exhibiting at
least 90% identity or which hybridizes with the sequences given in
ABQ81842 and ABQ81843. Also described is a polynucleotide (II) encoding a
fusion protein, comprising a sequence selected from 1097 sequences given
in ABP65258 to ABP65354 ligated in frame to a polynucleotide encoding a
heterologous polypeptide. (I) has antidiarrhetic and antibacterial
activities, and can be used as an inhibitor of salmonella. (I) (which is
a probe) is useful for the detection and/or identification of
Bifidobacterium longum in a biological sample. A carrier containing the
lactic acid bacterium Bifidobacterium longum NCC2705 (NCM I-2618) can be
used for preventing and/or treating diarrhoea brought about by pathogenic
bacteria and/or rotavirus. The carrier is a food composition selected
from milk, yogurt, curd, cheese, fermented milks, milk based fermented
products, ice-creams, fermented cereal based products, milk based
powders, infant formula, pet food or a pharmaceutical composition
selected from tablets, liquid bacterial suspensions, dried oral
supplement, wet oral supplement, dry tube feeding or wet tube feeding.
(I) is useful in DNA arrays or chips to carry out analysis of the
expression of the Bifidobacterium gene. ABQ81844 to ABQ81850 represent
Bifidobacterium related nucleotide sequences given in the Sequence
Listing from the present invention but not mentioned further within the
specification. N.B. the sequence data for this patent is not represented
in the printed specification but is based on sequence information
supplied by the European Patent Office

XX XX
Sequence 375 AA;

Query Match 45.1%; Score 46; DB 5; Length 375;
Best Local Similarity 47.4%; Pred. No. 50;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

1 PVKFFELPYGKQDPCASVLHYAQE 19

KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX
OS Hepatitis A virus.
OS Synthetic.
XX
PN WO200105824-A2.
XX
XX 25-JAN-2001.
PD
XX 14-JUL-2000; 2000WO-US019267.
PF
XX
XX 15-JUL-1999; 99US-0144412P.
PR
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX
XX Fields HA, Khudyakov YE;
PI
XX
XX WPI; 2001-112681/12.
DR
XX
XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.
PT
XX
XX
PS Claim 13; Page 94; 130pp; English.
XX

CC The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of HAV in a human or animal through the binding of
CC the peptide to an antibody, to detect acute phase infection by detecting
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity
XX

SQ Sequence 21 AA;

Query Match 44.1%; Score 45; DB 4; Length 21;
Best Local Similarity 100.0%; Pred.No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 QRLKYAQEE 20
| | | | | | | | | |
Db 1 QRLKYAQEE 9

Search completed: May 11, 2004, 13:37:25
Job time : 51 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:34:17 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLEVGKQRLKYAQQE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A COMB pep:*

2: /cgn2_6/prodata/2/iaa/5B COMB pep:*

3: /cgn2_6/prodata/2/iaa/6A COMB pep:*

4: /cgn2_6/prodata/2/iaa/6B COMB pep:*

5: /cgn2_6/prodata/2/iaa/6C COMB pep:*

6: /cgn2_6/prodata/2/iaa/6D COMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	1091	6	Patent No. 5516630-2
2	102	100.0	2227	3	US-08-475-886-2
3	102	100.0	2227	3	US-08-475-886-4
4	102	100.0	2227	3	US-08-475-886-6
5	102	100.0	2227	3	US-08-397-232-2
6	102	100.0	2227	3	US-08-397-232-4
7	102	100.0	2227	3	US-09-171-387-2
8	102	100.0	2227	4	US-09-653-499-2
9	102	100.0	2227	4	US-09-653-499-4
10	102	100.0	2227	4	US-09-653-499-6
11	102	100.0	2227	4	US-10-104-966-12
12	102	100.0	2227	4	US-10-135-988-2
13	102	100.0	2227	4	US-10-135-988-4
14	102	100.0	2227	4	US-10-135-988-6
15	94	92.2	839	1	US-08-087-015-2
16	47	46.1	249	4	US-09-328-352-4748
17	45	44.1	214	4	US-09-198-452A-108
18	43	42.2	50	4	US-09-621-976-4556
19	42	41.2	452	4	US-09-252-991A-28627
20	42	41.2	789	4	US-09-489-039A-9210
21	42	41.2	2190	4	US-09-252-991A-25754
22	41	40.2	60	4	US-09-621-976-4371
23	41	40.2	68	1	US-08-606-789-2
24	41	40.2	68	1	US-08-606-789-4
25	41	40.2	68	1	US-09-111-348-2
26	41	40.2	68	2	US-09-111-348-4
27	41	40.2	69	5	PCT-US95-06406A-5

5	PCT-US95-06406A-7	59	40.2	41	28
1	US-08-364-003B-2	147	40.2	41	29
2	US-08-959-865-3	147	40.2	41	30
3	US-08-842-234-2	147	40.2	41	31
4	US-09-671-317-488	147	40.2	41	32
5	US-09-134-000C-4845	189	40.2	41	33
6	US-09-252-991A-32086	534	40.2	41	34
7	US-09-252-991A-29088	794	40.2	41	35
8	US-09-252-991A-17763	796	40.2	41	36
9	US-09-439-261-21	182	39.2	40	37
10	US-09-227-613-20	182	39.2	40	38
11	US-09-080-643-2	189	39.2	40	39
12	US-09-080-643-4	189	39.2	40	40
13	US-09-439-261-20	219	39.2	40	41
14	US-09-227-613-19	219	39.2	40	42
15	US-09-439-261-13	287	39.2	40	43
16	US-09-227-613-14	287	39.2	40	44
17	US-09-439-261-14	288	39.2	40	45
18	US-09-439-261-16	288	39.2	40	46
19	US-09-439-261-18	288	39.2	40	47
20	US-09-227-613-15	288	39.2	40	48
21	US-09-439-261-42	347	39.2	40	49
22	US-09-227-613-40	347	39.2	40	50
23	US-09-439-261-41	360	39.2	40	51
24	US-09-227-613-39	360	39.2	40	52
25	US-09-439-261-11	444	39.2	40	53
26	US-09-227-613-12	444	39.2	40	54
27	US-09-439-261-43	444	39.2	40	55
28	US-09-227-613-42	444	39.2	40	56
29	US-09-048-888-3	444	39.2	40	57
30	US-09-439-261-39	445	39.2	40	58
31	US-09-439-261-45	445	39.2	40	59
32	US-09-252-991A-28079	487	39.2	40	60
33	US-09-198-452A-173	498	39.2	40	61
34	US-09-252-991A-243	586	39.2	40	62
35	US-09-370-838-67	764	39.2	40	63
36	US-08-157-005-3	1463	39.2	40	64
37	US-08-747-863-3	1463	39.2	40	65
38	US-09-565-864-3	1463	39.2	40	66
39	US-09-310-363C-2	384	38.7	39	67
40	US-08-477-831C-40	183	38.2	39	68
41	US-09-198-432A-723	244	38.2	39	69
42	US-09-134-000C-3690	347	38.2	39	70
43	US-09-310-363C-4	409	38.2	39	71
44	US-08-612-521-2	1408	38.2	39	72
45	US-08-056-200-94	1898	38.2	39	73
46	US-08-800-644-94	1898	38.2	39	74
47	US-08-529-055-32	170	37.7	38.5	75
48	US-09-328-352-6473	283	37.7	38.5	76
49	US-08-961-083-160	641	37.7	38.5	77
50	US-09-536-784-160	641	37.7	38.5	78
51	US-09-957-005-9	1007	37.7	38.5	79
52	US-08-714-741-32	8991	37.3	80	80
53	US-09-134-001C-3984	128	37.3	81	81
54	US-08-981-527A-11	129	37.3	82	82
55	US-09-489-039A-11444	192	37.3	83	83
56	US-09-134-001C-4162	211	37.3	84	84
57	US-08-836-236-8	233	37.3	85	85
58	US-08-719-758-2	256	37.3	86	86
59	US-09-119-827-2	256	37.3	87	87
60	US-09-252-991A-27270	261	37.3	88	88
61	US-09-543-681A-4956	391	37.3	89	89
62	US-09-252-991A-19593	584	37.3	90	90
63	US-09-252-991A-18292	1050	37.3	91	91
64	US-09-555-554-2	3072	37.3	92	92
65	US-09-413-814-93	3079	37.3	93	93
66	US-09-413-814-80	37.5	36.8	94	94
67	US-08-940-095-229	37.5	36.8	95	95
68	US-08-940-093-229	37.5	36.8	96	96
69	US-08-940-096-229	37.5	36.8	97	97
70	US-09-465-719-229	37.5	36.8	98	98
71	US-09-453-605-229	37.5	36.8	99	99
72	US-09-453-638-229	37.5	36.8	100	100

```

; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PR1
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
; US-08-475-886-4

Query Match      100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PYKELRLEVGKQRLKYAQEE 20
DB      799 PYKELRLEVGKQRLKYAQEE 818

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PR1
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; US-08-475-886-6

Query Match      100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PYKELRLEVGKQRLKYAQEE 20
DB      799 PYKELRLEVGKQRLKYAQEE 818

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H

```

```

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 255,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2
; LENGTH: 1091
5516630-2

Query Match          100.0%; Score 102; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PYKELRLEVKGKRLKYAQEE 20
        ||||||||||||||||
Db       1036 PYKELRLEVKGKRLKYAQEE 1055

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match          100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PYKELRLEVKGKRLKYAQEE 20
        ||||||||||||||||
Db       799 PYKELRLEVKGKRLKYAQEE 818

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A

```

APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 799 PYKELRLEVGKQRLKYAEE 818

RESULT 6
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US1
CURRENT APPLICATION NUMBER: US/08/397,232A
CURRENT FILING DATE: 1995-04-17
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: PCT/US93/08610
EARLIER FILING DATE: 1993-09-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 799 PYKELRLEVGKQRLKYAEE 818

RESULT 7
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
GENERAL INFORMATION:
APPLICANT: RAYCHAUDHURI, GOPA;
EMERSON, SUZANNE, U.;
PURCELL, ROBERT, H.
TITLE OF INVENTION: SIMIAN-HUMAN HAV
HAVING A CHIMERIC 2C PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feiler
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 102; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 799 PYKELRLEVGKQRLKYAEE 818

RESULT 8
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGVKQRLKYAQEE 20
Db 799 PYKELRLEVGVKQRLKYAQEE 818

```
Query Match      100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0
```



```
RESULT 13
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4
Query Match 100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELELVGKQRLKYAOEE 20
DB 799 PYKELELVGKQRLKYAOEE 818

RESULT 14
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
Query Match 100.0%; Score 102; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELELVGKQRLKYAOEE 20
DB 799 PYKELELVGKQRLKYAOEE 818

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135
; GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2
Query Match 92.2%; Score 94; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 2.9e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELELVGKQRLKYAOEE 20
DB 798 PYKELELVGKQRLKYAOEE 817

RESULT 16
US-09-328-352-4748
; Sequence 4748, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4748
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4748
Query Match 46.1%; Score 47; DB 4; Length 249;
Best Local Similarity 56.2%; Pred. No. 4;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```


; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.

US-09-252-991A-25754

Query Match 41.2%; Score 42; DB 4; Length 2190;

Best Local Similarity 47.4%; Pred. No. 2.6e-02; Mismatches 8; Indels 0; Gaps 0;

Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQRLKYAQ 19

DB 40 PIPQTRQEMGNPLSLVAQE 58

RESULT 22

US-09-621-976-4371

; Sequence 4371, Application US/09621976

; Patent No. 6639083

; GENERAL INFORMATION:

; APPLICANT: Dumas Maline Edwards, J.B.

; APPLICANT: Giordano, J.Y.

; TITLE OF INVENTION: ESTs and Encoded Human Proteins.

; FILE REFERENCE: GENSET.054PR2

; CURRENT APPLICATION NUMBER: US/09/621,976

; CURRENT FILING DATE: 2000-07-21

; NUMBER OF SEQ ID NOS: 19335

; SOFTWARE: Patent.pm

; SEQ ID NO 4371

; LENGTH: 60

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: 8

; OTHER INFORMATION: Xaa = Arg,Thr

; NAME/KEY: UNSURE

; LOCATION: 9

; OTHER INFORMATION: Xaa = His,Leu,Pro,Arg

; US-09-621-976-4371

Query Match

Best Local Similarity 40.2%; Score 41; DB 4; Length 60;

Matches 7; Conservative 6; Mismatches 3; Indels 0;

Matches 0; Gaps 0;

QY 3 KELRLEVGVKQRLKYAQ 18

DB 15 EQLKLEAGVRIKVSQ 30

RESULT 23

US-08-606-789-2

; Sequence 2, Application US/08606789

; Patent No. 5783418

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Stuart, Susan G.

; APPLICANT: Murry, Lynn E.

; APPLICANT: Guegler, Kark J.

; APPLICANT: Seilhamer, Jeffrey J.

; TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/606,789

; FILING DATE: Filed Herewith

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA: 08/440,743

; APPLICATION NUMBER: 08/440,743

; FILING DATE: May 5, 1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/320,011

; FILING DATE: October 5, 1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Luther, Barbara J.

; REGISTRATION NUMBER: 33,954

; REFERENCE/DOCKET NUMBER: PF-0055 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-852-0195

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 68 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; LIBRARY: PITUITARY

; CLONE: 112530

; US-08-606-789-2

Query Match 40.2%; Score 41; DB 1; Length 68;

Best Local Similarity 50.0%; Pred. No. 9.9;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 KELRLEVGVKQRLKYAQ 18

DB 15 QQLRLEAGLNVRKVSQ 30

RESULT 24

US-08-606-789-4

; Sequence 4, Application US/08606789

; Patent No. 5783418

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Stuart, Susan G.

; APPLICANT: Murry, Lynn E.

; APPLICANT: Guegler, Kark J.

; APPLICANT: Seilhamer, Jeffrey J.

; TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: U.S.

; ZIP: 94303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/606,789

; FILING DATE: Filed Herewith

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/440,743

; FILING DATE: May 5, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Luther, Barbara J.

; REGISTRATION NUMBER: 33,954

; REFERENCE/DOCKET NUMBER: PF-0055 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-852-0195

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 68 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; LIBRARY: PITUITARY

; CLONE: 112530

; US-08-606-789-2

/ FILING DATE: October 5, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Luther, Barbara J.
/ REGISTRATION NUMBER: 33,954
/ REFERENCE/DOCKET NUMBER: PF-0055 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-852-0195
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 68 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 204241
/ US-08-606-789-4

Query Match 40.2%; Score 41; DB 1; Length 68;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 KELRLEVGKQRLKYAQ 18
Db 15 QQLRLEAGLNVRKVSQ 30

RESULT 25
US-09-111-348-2
/ Sequence 2, Application US/09111348
/ Patent No. 5912130
/ GENERAL INFORMATION:
/ APPLICANT: Au-Young, Janice
/ APPLICANT: Stuart, Susan G.
/ APPLICANT: Murry, Lynn E.
/ APPLICANT: Guegler, Kark J.
/ APPLICANT: Seilhamer, Jeffrey J.
/ TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: U.S.
/ ZIP: 94303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/111,348
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/606,789
/ FILING DATE:
/ APPLICATION NUMBER: 08/440,743
/ FILING DATE: May 5, 1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/320,011
/ FILING DATE: October 5, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Luther, Barbara J.
/ REGISTRATION NUMBER: 33,954
/ REFERENCE/DOCKET NUMBER: PF-0055 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555

/ TELEFAX: 415-852-0195
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 68 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ IMMEDIATE SOURCE:
/ LIBRARY: PITUITARY
/ CLONE: 112530
/ US-09-111-348-2

Query Match 40.2%; Score 41; DB 2; Length 68;
Best Local Similarity 50.0%; Pred. No. 9.9;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 KELRLEVGKQRLKYAQ 18
Db 15 QQLRLEAGLNVRKVSQ 30

RESULT 26
US-09-111-348-4
/ Sequence 4, Application US/09111348
/ Patent No. 5912130
/ GENERAL INFORMATION:
/ APPLICANT: Au-Young, Janice
/ APPLICANT: Stuart, Susan G.
/ APPLICANT: Murry, Lynn E.
/ APPLICANT: Guegler, Kark J.
/ APPLICANT: Seilhamer, Jeffrey J.
/ TITLE OF INVENTION: HUMAN HOMOLOG OF THE RAT G PROTEIN
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: U.S.
/ ZIP: 94303
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq Version 1.5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/111,348
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/606,789
/ FILING DATE:
/ APPLICATION NUMBER: 08/440,743
/ FILING DATE: May 5, 1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/320,011
/ FILING DATE: October 5, 1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Luther, Barbara J.
/ REGISTRATION NUMBER: 33,954
/ REFERENCE/DOCKET NUMBER: PF-0055 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-852-0195
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 4:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 68 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 147 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-264-003B-2

Query Match 40.2%; Score 41; DB 1; Length 147;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LEVGKQLKY 16
|:|||||
Db 24 LOVGKARLKY 33

RESULT 30

US-08-959-865-3
Sequence 3, Application US/08959865
Patent No. 5915627

GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: MICROSOMAL GLUTATHIONE-S TRANSFERASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/959,865

FILING DATE: Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0411 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 147 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: COLNOT07

CLONE: 903729

US-08-959-865-3

Query Match 40.2%; Score 41; DB 2; Length 147;
Best Local Similarity 80.0%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 LEVGKQLKY 16

Db 24 LOVGKARLKY 33
|:|||||

Search completed: May 11, 2004, 13:42:20
Job time : 14.5 secs

OM protein - protein search, using sw model
Run on: May 11, 2004, 13:30:01 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-09-171-432A-40
Perfect score: 102
Sequence: 1 PYKELRLEVGKQLKYAQEE 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	55	2 P00433	genome polypotein
2	102	100.0	56	2 P00434	genome polypotein
3	102	100.0	56	2 P00428	genome polypotein
4	102	100.0	56	2 P00432	genome polypotein
5	102	100.0	56	2 P00430	genome polypotein
6	102	100.0	341	2 S04137	genome polypotein
7	102	100.0	852	1 GNNYHA	genome polypotein
8	102	100.0	1358	2 A03905	genome polypotein
9	102	100.0	2227	1 GNNYHM	genome polypotein
10	102	100.0	2227	1 GNNYHR	genome polypotein
11	102	100.0	2227	1 GNNYMK	genome polypotein
12	102	100.0	2227	1 GNNYHE	genome polypotein
13	100	98.0	56	2 P00427	genome polypotein
14	99	97.1	56	2 P00429	genome polypotein
15	95	93.1	56	2 P00436	genome polypotein
16	95	93.1	2230	1 GNNYSA	genome polypotein
17	94	92.2	55	2 P00435	genome polypotein
18	94	92.2	56	2 P00437	genome polypotein
19	94	92.2	839	1 GNNYS2	genome polypotein
20	93	91.2	56	2 P00431	genome polypotein
21	62	60.8	319	2 JH0135	genome polypotein
22	46	45.1	219	2 T45997	hypothetical prote
23	46	45.1	285	2 T29804	hypothetical prote
24	45	44.1	209	2 D72121	ATP synthase chain
25	45	44.1	209	2 B86502	ATP synthase subun
26	44	43.1	179	2 AI0819	probable exported
27	44	43.1	290	2 C71701	branched-chain ami
28	44	43.1	776	2 T05938	hypothetical prote
29	43	42.2	207	2 H85609	hypothetical prote

30	43	42.2	207	2	A99801	hypothetical prote
31	43	42.2	208	1	F70485	uracil phosphoribo
32	43	42.2	465	2	T25441	hypothetical prote
33	43	42.2	509	2	G81929	probable iron-upta
34	43	42.2	534	2	J05096	transposase - fung
35	43	42.2	3587	2	T31075	tyrocidine synthet
36	42.5	41.7	179	2	H91049	probable membrane
37	42.5	41.7	179	2	E85894	hypothetical prote
38	42.5	41.7	179	2	A65027	ATP-dependent heli
39	42.5	41.7	722	2	A02222	hypothetical prote
40	42	41.2	290	2	B97774	probable cysteine
41	42	41.2	365	2	F86413	probable porin PA0
42	42	41.2	452	2	B83623	flagellar biosynth
43	42	41.2	695	2	AH2647	flagellar biosynth
44	42	41.2	723	2	G97429	ribose/galactose A
45	42	41.2	736	2	B82944	DNA topoisomerase
46	42	41.2	752	2	A45582	DNA topoisomerase
47	42	41.2	752	2	A65089	DNA topoisomerase
48	42	41.2	752	2	G91116	DNA topoisomerase
49	42	41.2	752	2	G85961	topoisomerase IV c
50	42	41.2	752	2	AF0888	hypothetical prote
51	42	41.2	826	2	H84683	valine-tRNA ligase
52	42	41.2	939	2	H71532	protein T6D22.8 (i
53	42	41.2	990	2	A86215	hypothetical prote
54	42	41.2	4131	2	T21085	GTP-binding regula
55	41	40.2	68	2	B42243	GTP-binding regula
56	41	40.2	68	2	I39158	conserved hypotet
57	41	40.2	151	2	A72409	hypothetical prote
58	41	40.2	384	2	E84188	aminopeptidase II
59	41	40.2	412	2	E70108	probable atpH prot
60	41	40.2	446	2	G70774	hypothetical prote
61	41	40.2	515	2	H75579	ATP-dependent RNA
62	41	40.2	585	2	G96995	hypothetical prote
63	41	40.2	635	2	D84920	hypothetical prote
64	41	40.2	656	2	E75468	hypothetical prote
65	41	40.2	735	2	E83540	conserved hypotet
66	41	40.2	777	2	B95991	heat shock atp-dep
67	41	40.2	842	2	G90576	ClpB protein NMA16
68	41	40.2	859	2	F81863	ClpB protein NMB14
69	41	40.2	859	2	F81078	RNA-directed RNA p
70	41	40.2	1116	2	T30828	hypothetical prote
71	41	40.2	1206	2	D90085	molybdenum-pterin-
72	40	39.2	69	2	I64119	hypothetical prote
73	40	39.2	178	2	F87408	peptidyl-tRNA hyd
74	40	39.2	189	2	E95000	aminoacyl-tRNA hyd
75	40	39.2	189	2	E97872	uridine kinase (EC
76	40	39.2	213	2	AH0185	DNA-binding respon
77	40	39.2	218	2	C82378	hypothetical prote
78	40	39.2	241	2	A71020	hypothetical prote
79	40	39.2	353	2	F69258	iron-sulfur bindin
80	40	39.2	366	2	B69113	translation elonga
81	40	39.2	456	2	S11665	BFR1 protein - Yea
82	40	39.2	470	2	S47887	hypothetical prote
83	40	39.2	485	2	E86506	hypothetical prote
84	40	39.2	485	2	E72115	hypothetical prote
85	40	39.2	485	2	A81555	hypothetical prote
86	40	39.2	501	2	C71948	RNA-directed RNA p
87	40	39.2	533	2	A45392	flagellar biosynth
88	40	39.2	583	2	A83530	methionine-tRNA li
89	40	39.2	723	2	D71091	autoantigen NOR-90
90	40	39.2	727	2	S18193	ribosomal transcri
91	40	39.2	727	2	JCS113	UBF transcription
92	40	39.2	727	2	B40439	transcription fact
93	40	39.2	764	2	S03318	ribosomal transcri
94	40	39.2	764	2	JCS112	UBF transcription
95	40	39.2	764	2	A40439	transcription fact
96	40	39.2	765	2	S22314	endopeptidase Ia (
97	40	39.2	795	1	S73830	outer arm dynein i
98	40	39.2	837	2	T02761	valyl-tRNA synthet
99	40	39.2	939	2	H81686	ori 1b protein - L
100	40	39.2	1463	2	A36861	

ALIGNMENTS

RESULT 1

PQ0433
genome polyprotein - human hepatitis A virus (strain PA21) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0433
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0433
A:Molecule type: mRNA
A:Residues: 1-55 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 35 PYKELRLEVGKQRLKYAQEE 54

RESULT 2

PQ0434
genome polyprotein - human hepatitis A virus (strain KPH) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0434
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0434
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 3

PQ0428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0428
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 4

PQ0432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0432
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0432
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 5

PQ0430
genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0430
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0430
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 6

S04137
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
C:Species: human hepatitis A virus
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A:Reference number: S04137; MUID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>

Query Match 100.0%; Score 102; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20

Db 36 PYKELRLEVGKQRLKYAQEE 55

A;Cross-references: EMBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576

C;Genetics:

A;Gene: VP1

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; polyprotein

F;2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 100.0%; Score 102; DB 2; Length 341;

Best Local Similarity 100.0%; Pred. No. 1.3e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

Qy 1 PYKELRLEVGKQRLKYAOEE 20

Db 303 PYKELRLEVGKQRLKYAOEE 322

RESULT 7

GNVYHA

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)

N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C;Species: human hepatitis A virus

A;Note: host Homo sapiens (man)

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C;Accession: A03904

R;Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.

J. Virol. 54, 247-255, 1985

A;Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A;Reference number: A03904; MUID:85185648; PMID:2985793

A;Accession: A03904

A;Molecule type: genomic RNA

A;Residues: 1-852 <LIN>

A;Cross-references: EMBL:M10033; NID:G329592; PIDN:AAA45470.1; PID:G329593

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; polyprotein

F;1-245/Product: coat protein 1A #status predicted <C1A>

F;246-491/Product: coat protein 1B #status predicted <C1B>

F;492-836/Product: coat protein 1C #status predicted <C1C>

F;837-852/Product: coat protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 102; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. No. 3.7e-08; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

Qy 1 PYKELRLEVGKQRLKYAOEE 20

Db 799 PYKELRLEVGKQRLKYAOEE 818

RESULT 8

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein

C;Species: human hepatitis A virus

C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996

C;Accession: A03905

R;Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinston

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A;Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA

A;Reference number: A03905; MUID:85166289; PMID:2984684

A;Accession: A03905

A;Molecule type: genomic RNA

A;Residues: 1-1358 <BAR>

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; polyprotein

F;1-245/Product: coat protein 1A #status predicted <C1A>

F;246-491/Product: coat protein 1B #status predicted <C1B>

F;492-836/Product: coat protein 1C #status predicted <C1C>

F;837-854/Product: coat protein 2A (fragment) #status predicted <C2A>

F;855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 102; DB 2; Length 1358;

Best Local Similarity 100.0%; Pred. No. 6.1e-09; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

Qy 1 PYKELRLEVGKQRLKYAOEE 20

Db 799 PYKELRLEVGKQRLKYAOEE 818

RESULT 9

GNVYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core prot

B; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C;Species: human hepatitis A virus

A;Note: host Homo sapiens (man)

C;Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C;Accession: A25981

R;Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A;Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with dif

A;Reference number: A25981; MUID:87061253; PMID:3023706

A;Accession: A25981

A;Molecule type: genomic RNA

A;Residues: 1-2227 <COH>

A;Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransf

F;1-23/Product: coat protein 1A #status predicted <VP1>

F;24-245/Product: coat protein 1B #status predicted <VP2>

F;246-491/Product: coat protein 1C #status predicted <VP3>

F;492-791/Product: coat protein 1D #status predicted <VP4>

F;792-980/Product: coat protein 2A #status predicted <C2A>

F;981-1087/Product: core protein 2B #status predicted <C2B>

F;1088-1422/Product: core protein 2C #status predicted <C2C>

F;1423-1496/Product: protein 3A #status predicted <C3A>

F;1497-1519/Product: protein 3B #status predicted <C3B>

F;1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F;1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 102; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.1e-07; Mismatches 20; Conservative 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

Qy 1 PYKELRLEVGKQRLKYAOEE 20

Db 799 PYKELRLEVGKQRLKYAOEE 818

RESULT 10

GNVYHR

genome polyprotein - human hepatitis A virus

N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core prot

NA polymerase (EC 2.7.7.48), protein 3D

C;Species: human hepatitis A virus

A;Note: host Homo sapiens (man)

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999

C;Accession: A03903

R;Najarian, R.; Caput, D.; Ges, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest,

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A;Title: Primary structure and gene organization of human hepatitis A virus.

A;Reference number: A03903; MUID:85190549; PMID:2986127

A;Accession: A03903

A;Molecule type: genomic RNA

A;Residues: 1-2227 <NAJ>

A;Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597

C;Superfamily: hepatitis A virus genome polyprotein

C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransf

F;1-245/Product: coat protein 1A #status predicted <C1A>

F;246-491/Product: coat protein 1B #status predicted <C1B>

F;492-836/Product: coat protein 1C #status predicted <C1C>

F;837-980/Product: core protein 2A #status predicted <C2A>

F;981-1076/Product: core protein 2B #status predicted <C2B>

F;1077-1422/Product: core protein 2C #status predicted <C2C>

F;1423-1484/Product: protein 3A #status predicted <C3A>

F;1485-1507/Product: protein 3B #status predicted <C3B>

F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>
Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 799 PYKELRLEVGVKQRLKYAQEE 818
RESULT 11
GNNYMK
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 2A; core protein 3D
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text_change 16-Jul-1999
C:Accession: A94149; A25914; A94508
R:Coher, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R.
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
A:Reference number: A94149; MUID:87175701; PMID:3031686
A:Accession: A94149
A:Status: nucleic acid sequence not shown
A:Molecule type: genomic RNA
A:Residues: 1-2227 <CDS>
A:Cross-references: EMBL:M16632; NID:G329594; PID:AAA45471.1; PID:G329595
A:Note: submitted to GenBank, August 1987
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F:1-245/Product: coat protein 1A #status predicted <P1A>
F:246-491/Product: coat protein 1B #status predicted <P1B>
F:492-836/Product: coat protein 1C #status predicted <P1C>
F:837-980/Product: coat protein 2A #status predicted <P2A>
F:981-1076/Product: core protein 2B #status predicted <P2B>
F:1077-1422/Product: core protein 2C #status predicted <P2C>
F:1423-1484/Product: protein 3A #status predicted <P3A>
F:1485-1507/Product: protein 3B #status predicted <P3B>
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>
Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 799 PYKELRLEVGVKQRLKYAQEE 818
RESULT 12
GNNYHB
genome polyprotein - human hepatitis A virus (strain MBB)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 3D
VPg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text_change 10-May-1996
C:Accession: J50303
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Kiehn, R.; Wimmer, B.; Deinhardt, V.
Virus Res. 8, 153-171, 1997
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolated from a patient with acute hepatitis A)
A:Reference number: J50303; MUID:88045071; PMID:2823500
A:Accession: J50303
A:Molecule type: genomic RNA
A:Residues: 1-2227 <PAU>
A:Cross-references: EMBL:M20273
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrophobic protein
F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-246/Product: coat protein 1B #status predicted <VP2>
F:247-494/Product: coat protein 1C #status predicted <VP3>
F:492-836/Product: coat protein 1D #status predicted <VP1>
F:837-980/Product: core protein 2A #status predicted <P2A>
F:981-1108/Product: core protein 2B #status predicted <P2B>
F:1109-1438/Product: core protein 2C #status predicted <P2C>
F:1439-1496/Product: core protein 3A #status predicted <P3A>
F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>
Query Match 100.0%; Score 102; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 799 PYKELRLEVGVKQRLKYAQEE 818
RESULT 13
PQ0427
genome polyprotein - human hepatitis A virus (strain Ep-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J.
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Comment: This protein is from the VP1/2A Junction region.
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 98.0%; Score 100; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 3.8e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 14
PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J.
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A:Reference number: PQ0429; MUID:92300330; PMID:1318940
A:Accession: PQ0429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
Query Match 97.1%; Score 99; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 5.5e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 15

P00436
genome polyprotein - human hepatitis A virus (strain AGM27) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00436
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00436
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 93.1%; Score 95; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 2.4e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||:|||||:|||||
Db 36 PYKELRMEVGKQRLKYAMEE 55
|||:|||||:|||||

RESULT 16
GNVYS2
genome polyprotein - simian hepatitis A virus (strain AGM-27)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
C:Species: simian hepatitis A virus
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C:Accession: A30470; S04885; S03965
R:Tsarev, S.A.
submitted to JTPID, April 1991
A:Reference number: A30470
A:Accession: A30470
A:Molecule type: genomic RNA
A:Residues: 1-2230 <TSA>
A:Cross-references: GB:D00924; NID:9222597; PIDN:BA00766.1; PID:9222598
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.
J. Gen. Virol. 72, 1677-1683, 1991
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an
A:Reference number: J01080; MUID:91311420; PMID:1645901
A:Contents: annotation
A>Note: neither amino acid nor nucleotide sequence is given
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tearev, S.A.; Sverdlov, E.D.; Chizhik
submitted to the EMBL Data Library, May 1989
A:Reference number: S04885
A:Accession: S04885
A:Molecule type: genomic RNA
A:Residues: 1750-2164 <BAL1>
A:Cross-references: EMBL:X15461; NID:961971; PIDN:CAA33490.1; PID:930268
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tearev, S.A.; Sverdlov, E.D.; Chizhik
FEBS Lett. 247, 423-428, 1989
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A:Reference number: S03965; MUID:89232168; PMID:2541023
A:Accession: S03965
A:Molecule type: genomic RNA
A:Residues: 1960-2164 <BAL2>
A:Cross-references: EMBL:X15461
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-27/Product: coat protein 1A #status predicted <C1A>
F:28-249/Product: coat protein 1B #status predicted <C1B>
F:250-495/Product: coat protein 1C #status predicted <C1C>
F:496-795/Product: coat protein 1D #status predicted <C1D>
F:796-984/Product: core protein 2A #status predicted <C2A>
F:985-1091/Product: core protein 2B #status predicted <C2B>
F:1092-1426/Product: core protein 2C #status predicted <C2C>
F:1427-1498/Product: protein 3A #status predicted <P3A>
F:1499-1521/Product: protein 3B #status predicted <P3B>
F:1522-1741/Product: protein 3C #status predicted <P3C>
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 93.1%; Score 95; DB 1; Length 2230;
Best Local Similarity 90.0%; Pred. No. 1.4e-06;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||:|||||:|||||
Db 803 PYKELRMEVGKQRLKYAMEE 822
|||:|||||:|||||

RESULT 17
P00435
genome polyprotein - human hepatitis A virus (strain Cyl145) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00435
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00435
A:Molecule type: mRNA
A:Residues: 1-55 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 92.2%; Score 94; DB 2; Length 55;
Best Local Similarity 90.0%; Pred. No. 3.4e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||:|||||:|||||
Db 35 PYKELRLEVGKQRLKYAMEE 54
|||:|||||:|||||

RESULT 18
P00437
genome polyprotein - human hepatitis A virus (strain JM55) (fragment)
C:Species: human hepatitis A virus
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00437
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00437
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 92.2%; Score 94; DB 2; Length 56;
Best Local Similarity 90.0%; Pred. No. 3.5e-08;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
|||:|||||:|||||
Db 36 PYKELRMEVGKQRLKYAMEE 55
|||:|||||:|||||

RESULT 19
GNVYS2
genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein
C:Species: simian hepatitis A virus
A>Note: host Macaca fascicularis (cynomolgus macaque)
C>Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 16-Jul-1999
C:Accession: J01180
R:Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
J. Gen. Virol. 72, 1685-1689, 1991
A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus mac

A:Reference number: JQ1180; MUID:913111421; PMID:1649902
A:Accession: JQ1180
A:Molecule type: Genomic RNA
A:Residues: 1-839 <NAI>
A:Cross-references: GB:M59286; NID:G329599; PID:AAA45473.1; PID:G555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.2%; Score 94; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 6.9e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQRLKYAQEE 20
|||||
DB 798 PYKELRLEVGVKQRLKYAQEE 817

RESULT 20
PQ0431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
A:Accession: PQ0431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0431
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 91.2%; Score 93; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 5e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQRLKYAQEE 20
|||||
DB 36 PYKELRLEVGVKQRLKYAQEE 55

RESULT 21
JH0135
genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994
A:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
Virus Res. 8, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus.
A:Reference number: JH0135; MUID:88129044; PMID:2829458
A:Accession: JH0135
A:Molecule type: genomic RNA
A:Residues: 1-319 <ROS>
A:Cross-references: GB:M22821
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-300/Product: coat protein 1B #status predicted <CPD>
F:301-319/Product: core protein 2A (fragment) #status predicted <C2h>
F:237/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.8%; Score 62; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Reference number: JQ1180; MUID:913111421; PMID:1649902
A:Accession: JQ1180
A:Molecule type: Genomic RNA
A:Residues: 1-839 <NAI>
A:Cross-references: GB:M59286; NID:G329599; PID:AAA45473.1; PID:G555083
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-23/Product: coat protein 1A #status predicted <VP0>
F:24-245/Product: coat protein 1B #status predicted <VP3>
F:246-491/Product: coat protein 1C #status predicted <VP1>
F:492-839/Product: core protein 2A (fragment) #status predicted <P2P>
F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 92.2%; Score 94; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 6.9e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQRLKYAQEE 20
|||||
DB 798 PYKELRLEVGVKQRLKYAQEE 817

RESULT 20
PQ0431
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
A:Accession: PQ0431
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel
J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0431
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 91.2%; Score 93; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 5e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQRLKYAQEE 20
|||||
DB 36 PYKELRLEVGVKQRLKYAQEE 55

RESULT 21
JH0135
genome polyprotein - human hepatitis A virus (strain MS-1) (fragment)
N:Contains: amino end of core protein 2A; coat protein 1D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994
A:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
Virus Res. 8, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus.
A:Reference number: JH0135; MUID:88129044; PMID:2829458
A:Accession: JH0135
A:Molecule type: genomic RNA
A:Residues: 1-319 <ROS>
A:Cross-references: GB:M22821
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; glycoprotein; polyprotein
F:1-300/Product: coat protein 1B #status predicted <CPD>
F:301-319/Product: core protein 2A (fragment) #status predicted <C2h>
F:237/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.8%; Score 62; DB 2; Length 319;
Best Local Similarity 100.0%; Pred. No. 0.032;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGVKQ 12
|||||
DB 308 PYKELRLEVGVKQ 319

RESULT 22
T45997
hypothetical protein F9D24.280 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 02-Sep-2000
A:Accession: T45997
R:D'Angelio, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.;
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23011
A:Accession: T45997
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-219 <DAN>
A:Cross-references: EMBL:AL137081
A:Experimental source: cultivar Columbia; BAC clone F9D24
C:Genetics:
A:Map position: 3
A:Introns: 85/3
A:Note: F9D24.280
C:Superfamily: Arabidopsis thaliana hypothetical protein F9D24.210

Query Match 45.1%; Score 46; DB 2; Length 219;
Best Local Similarity 50.0%; Pred. No. 7.9;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 KELRLEVGVKQRLKYAQEE 20
|||
DB 167 KKKVETGKARLQRAEE 184

RESULT 23
T23804
hypothetical protein C06E4.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
A:Accession: T23804
R:Du, Z.; Gattung, S.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C06E4.
A:Reference number: Z20688
A:Accession: T23804
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-285 <DUZ>
A:Cross-references: EMBL:U41277; PIDN:AAA82479.1; CESP:C06E4.8
C:Genetics:
A:Gene: CESP:C06E4.8
A:Introns: 31/2; 127/3; 153/3; 212/3
C:Superfamily: Caenorhabditis elegans hypothetical protein C06E4.8

Query Match 45.1%; Score 46; DB 2; Length 285;
Best Local Similarity 60.0%; Pred. No. 11;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 ELRLEVGVKQRLKYAQ 18
|||||
DB 42 ELRRQIGKQKHYEQ 56

RESULT 24
D72121
ATP synthase chain D - Chlamydomonas reinhardtii (strain CWL029)
C:Species: Chlamydomonas reinhardtii, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 05-May-2000
A:Accession: D72121
R:Kallman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: D72121
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209 <ARN>
A:Cross-references: GB:AE001594; GB:AE001363; NID:g4376341; PIDN:AA018243.1; PID:g437634
A:Experimental source: strain CWL029
C:Genetics:
A:Gene: atpD
C:Superfamily: H+-transporting ATPase chain D

Query Match 44.1%; Score 45; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 KKLRELVGKQRLKYAQEE 20
|||: |||: |||: |||:
Db 133 KKVMAEVSKERLKILEE 150

RESULT 25
B86502
ATP synthase subunit D [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Aug-2002
A:Accession: B86502
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: B86502
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209 <STO>
A:Cross-references: GB:BA000008; NID:g8978463; PIDN:BA98300.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: atpD
C:Superfamily: H+-transporting ATPase chain D

Query Match 44.1%; Score 45; DB 2; Length 209;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 KKLRELVGKQRLKYAQEE 20
|||: |||: |||: |||:
Db 133 KKVMAEVSKERLKILEE 150

RESULT 26
A10819
probable exported protein STV2748 [imported] - Salmonella enterica subsp. enterica serov
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
A:Accession: A10819
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: A10819
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02709.1; PID:g16503724; GSPDB:GN00176
C:Genetics:
A:Gene: STV2748

Query Match 43.1%; Score 44; DB 2; Length 179;
Best Local Similarity 52.9%; Pred. No. 67;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 ELRLVGVKQRLKYAQEE 20
|||: |||: |||: |||:
Db 569 ELRLVGVKQRLKYAQEE 595

RESULT 29
H85609
hypothetical protein Z1195 [imported] - Escherichia coli (strain O157:H7, substrain EDL95:
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A:Accession: H85609; B85661
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: C71701
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14885.1; PID:g386098;
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: ilvE; RP428
C:Superfamily: branched-chain-amino-acid aminotransferase

Query Match 43.1%; Score 44; DB 2; Length 290;
Best Local Similarity 61.1%; Pred. No. 23;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KKLRELVGKQRLKYAQEE 20
|||: |||: |||: |||:
Db 225 KSLCLEVSKERLKLAQIE 242

RESULT 28
T09938
hypothetical protein T16L4.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
A:Accession: T09938
R:Bavan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16897
A:Accession: T09938
A:Molecule type: DNA
A:Residues: 1-776 <BEV>
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.260
A:Experimental source: cultivar Columbia; BAC clone T16L4
C:Genetics:
A:Gene: ATSP:T16L4.260
A:Map position: 4
A:Introns: 308/3; 419/1; 462/3; 482/1; 574/3; 627/1; 724/3; 744/3
Query Match 43.1%; Score 44; DB 2; Length 776;
Best Local Similarity 52.9%; Pred. No. 67;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 ELRLVGVKQRLKYAQEE 20
|||: |||: |||: |||:
Db 569 ELRLVGVKQRLKYAQEE 595

RESULT 29
H85609
hypothetical protein Z1195 [imported] - Escherichia coli (strain O157:H7, substrain EDL95:
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A:Accession: H85609; B85661
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Best Local Similarity 60.0%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RLEVGKQRLKYAQEE 20
|||: |||: |||: |||:
Db 107 RLEIQKQLALAREE 121

RESULT 27
C71701
branched-chain amino acid aminotransferase (ilvE) RP428 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
A:Accession: C71701
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: C71701
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <AND>
A:Cross-references: GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA14885.1; PID:g386098;
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: ilvE; RP428
C:Superfamily: branched-chain-amino-acid aminotransferase

Query Match 43.1%; Score 44; DB 2; Length 290;
Best Local Similarity 61.1%; Pred. No. 23;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 KKLRELVGKQRLKYAQEE 20
|||: |||: |||: |||:
Db 225 KSLCLEVSKERLKLAQIE 242

RESULT 28
T09938
hypothetical protein T16L4.260 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 15-Oct-1999
A:Accession: T09938
R:Bavan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16897
A:Accession: T09938
A:Molecule type: DNA
A:Residues: 1-776 <BEV>
A:Cross-references: EMBL:AL079344; GSPDB:GN00062; ATSP:T16L4.260
A:Experimental source: cultivar Columbia; BAC clone T16L4
C:Genetics:
A:Gene: ATSP:T16L4.260
A:Map position: 4
A:Introns: 308/3; 419/1; 462/3; 482/1; 574/3; 627/1; 724/3; 744/3
Query Match 43.1%; Score 44; DB 2; Length 776;
Best Local Similarity 52.9%; Pred. No. 67;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 4 ELRLVGVKQRLKYAQEE 20
|||: |||: |||: |||:
Db 569 ELRLVGVKQRLKYAQEE 595

RESULT 29
H85609
hypothetical protein Z1195 [imported] - Escherichia coli (strain O157:H7, substrain EDL95:
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
A:Accession: H85609; B85661
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:24:26 / Search time 6.75 Seconds
(without alignment)

154.282 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102

Sequence: 1 PYKELRLVGVQRLKYAQEE 20

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	100.0	341	1	POLG_HPAVL
2	102	100.0	808	1	POLG_HPAVC
3	102	100.0	852	1	POLG_HPAVC
4	102	100.0	2226	1	POLG_HPAV2
5	102	100.0	2226	1	POLG_HPAV4
6	102	100.0	2226	1	POLG_HPAV8
7	102	100.0	2227	1	POLG_HPAVH
8	102	100.0	2227	1	POLG_HPAVL
9	102	100.0	2227	1	POLG_HPAVM
10	95	93.1	2230	1	POLG_HPAVS
11	94	92.2	839	1	POLG_HPAVT
12	45	44.1	209	1	VATD_CHLPN
13	44	43.1	290	1	ILVE_RICPR
14	44	43.1	583	1	FOJO_DROME
15	44	43.1	720	1	CANC_MOUSE
16	43.5	42.6	1819	1	GCP6_MOUSE
17	43	42.2	208	1	UPP_AQUAE
18	43	42.2	480	1	DNAH_RHIME
19	43	42.2	3587	1	TYCB_BREPA
20	42.5	41.7	179	1	YFGI_ECOLI
21	42	41.2	290	1	ILVE_RICGN
22	42	41.2	752	1	PARC_ECOLI
23	42	41.2	752	1	PARC_SALTY
24	42	41.2	939	1	SVY_CHLTR
25	41	40.2	68	1	BGGS_HUMAN
26	41	40.2	68	1	BGGA_HUMAN
27	41	40.2	147	1	GST2_HUMAN
28	41	40.2	318	1	Y234_AQUAE
29	41	40.2	446	1	ATPD_MYCTU
30	40	39.2	69	1	MOP_HAEIN
31	40	39.2	213	1	PTH_STRPN
32	40	39.2	189	1	URK_YERPE
33	40	39.2	353	1	Y070_ARCFU
					P13672 hepatitis a
					Q02381 hepatitis a
					P06442 hepatitis a
					P26580 hepatitis a
					P26581 hepatitis a
					P26582 hepatitis a
					P06617 hepatitis a
					P06441 hepatitis a
					P13901 hepatitis a
					P14553 simian hepa
					P31788 simian hepa
					Q92991 chlamydia p
					O05970 rickettsia
					P54360 drosophila
					Q96r56 mus musculus
					O67914 aquifex aeo
					P35890 rhizobium m
					O30408 b tyrocidin
					P76573 escherichia
					Q92126 rickettsia
					P20082 escherichia
					P26973 salmonella
					O84304 chlamydia t
					P30670 homo sapien
					P50151 homo sapien
					Q99735 homo sapien
					O66423 aquifex aeo
					Q10594 mycobacteri
					P45183 haemophilus
					Q97cd1 streptococc
					Q8zfz9 versinia pe
					O30166 archaeoglob

RESULT 1

ALIGNMENTS

34	40	39.2	456	1	EF1A_DICDI
35	40	39.2	470	1	BFRI_YEAST
36	40	39.2	571	1	ILVD_STRMU
37	40	39.2	685	1	STM1_HUMAN
38	40	39.2	685	1	STM1_MOUSE
39	40	39.2	723	1	SYM_PRRHO
40	40	39.2	733	1	ERG7_RAT
41	40	39.2	764	1	UBF1_HUMAN
42	40	39.2	764	1	UBF1_RAT
43	40	39.2	765	1	UBF1_MOUSE
44	40	39.2	795	1	LON_MYCPN
45	40	39.2	939	1	SYV_CHLMU
46	40	39.2	3859	1	RPOA_LELV
47	39.5	38.7	976	1	SN21_HUMAN
48	39.5	38.7	1009	1	SN2L_CAEEL
49	39.5	38.7	1812	1	BRC1_MOUSE
50	39	38.2	46	1	DIUH_LOCOMI
51	39	38.2	283	1	Y191_CIOAB
52	39	38.2	362	1	ILVE_STRCO
53	39	38.2	402	1	APL3_HUMAN
54	39	38.2	424	1	EXON_NFVOP
55	39	38.2	495	1	THIC_LEPIN
56	39	38.2	496	1	DNAH_BRUME
57	39	38.2	508	1	DNAH_BRUSU
58	39	38.2	508	1	V56K_PLRV1
59	39	38.2	508	1	V56K_PLRVW
60	39	38.2	553	1	PABP_SCHPO
61	39	38.2	1427	1	REST_HUMAN
62	39	38.2	1898	1	TRHY_HUMAN
63	39	38.2	1901	1	YCFL_TOBAC
64	38.5	37.7	416	1	RPSD_MICAE
65	38.5	37.7	641	1	DNAK_METSS
66	38	37.3	186	1	TNR2_ECOLI
67	38	37.3	226	1	PMT_MOUSE
68	38	37.3	239	1	PTH_RHIME
69	38	37.3	264	1	RPOD_SULAC
70	38	37.3	275	1	NAD2_ECOLI
71	38	37.3	289	1	SGCD_MESAU
72	38	37.3	289	1	SGCD_MOUSE
73	38	37.3	290	1	SGCD_HUMAN
74	38	37.3	298	1	YN05_YEAST
75	38	37.3	365	1	RECF_CHLMU
76	38	37.3	406	1	YG43_YEAST
77	38	37.3	420	1	YAGN_SCHPO
78	38	37.3	480	1	ILSB_CUCMA
79	38	37.3	490	1	PIT_BUCAP
80	38	37.3	493	1	ALGE_PSESM
81	38	37.3	493	1	PIT_BUCAP
82	38	37.3	799	1	SCA_DROME
83	38	37.3	802	1	OPHI_MOUSE
84	38	37.3	1050	1	BU1B_HUMAN
85	38	37.3	1052	1	BU1B_MOUSE
86	38	37.3	1966	1	MYSE_CAEEL
87	37.5	36.8	495	1	TRME_TREPA
88	37.5	36.8	522	1	GAG_HVZG1
89	37.5	36.8	1139	1	NGAP_HUMAN
90	37	36.3	42	1	BGG7_MOUSE
91	37	36.3	68	1	BGG7_BOVIN
92	37	36.3	68	1	BGG7_HUMAN
93	37	36.3	69	1	BGG7_RAT
94	37	36.3	85	1	RS16_PSESM
95	37	36.3	103	1	RL23_AQUAE
96	37	36.3	118	1	RNPA_VIBPA
97	37	36.3	118	1	RNPA_VIBVU
98	37	36.3	122	1	YP48_METJA
99	37	36.3	140	1	Y517_METJA
100	37	36.3	172	1	YC46_PYRAB
					P18624 dictyosteli
					P38934 saccharomyc
					Q8dt77 streptococc
					Q13586 homo sapien
					P70302 mus musculu
					O58721 pyrococcus
					P48450 rattus norv
					P25977 rattus norv
					P25976 mus musculu
					P78025 mycoplasma
					Q9p891 chlamydia m
					Q04561 lelystad vi
					P28370 homo sapien
					P41877 caenorhabdi
					P48754 mus musculu
					P23465 locusta mig
					Q97nk5 clostridium
					O86505 streptomyce
					O95236 homo sapien
					P24081 orgvia pseu
					Q8f795 leptospira
					Q8y6d5 brucella me
					Q8g3e7 brucella su
					P17525 potato leaf
					P11626 potato leaf
					P31209 schizosacch
					P30822 homo sapien
					Q07282 homo sapien
					P12222 nicotiana t
					P52322 microcystis
					Q92f66 methylovoru
					P04130 escherichia
					P23506 mus musculu
					Q92h67 rhizobium m
					P39471 sulfolobus
					P18843 escherichia
					P97281 mesocricetu
					P82347 mus musculu
					Q92829 homo sapien
					P53843 saccharomyc
					Q9pkw5 chlamydia m
					P53298 saccharomyc
					Q09873 schizosacch
					P13744 cucurbita m
					P57647 buchnera ap
					Q887q2 pseudomonas
					Q8k903 buchnera ap
					P21520 drosophila
					Q99131 mus musculu
					O60566 homo sapien
					Q921s0 mus musculu
					P02566 caenorhabdi
					O83561 treponema p
					P18041 human immun
					Q9u1f2 homo sapien
					O61016 mus musculu
					P30671 bos taurus
					O60262 homo sapien
					P43425 rattus norv
					Q886v2 pseudomonas
					O66433 aquifex aeo
					Q87tr4 vibrio para
					Q8ddi3 vuln para
					O58943 methanococc
					Q57937 methanococc
					Q9uzas pyrococcus

```

POLG HPAV1
ID POLG HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JUN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP3 (1C); Coat protein VP1 (1D); Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14666; CAA32794.1; -
DR PIR; S04137;
DR InterPro; IPR008975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN <1 1 COAT PROTEIN VP3.
FT CHAIN 2 340 COAT PROTEIN VP1.
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
FT SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
SQ
Query Match 100.0%; Score 102; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGKQRLKYAQEE 20
Db 303 PYKELRLEVGKQRLKYAQEE 322

RESULT 2
POLG HPAVG
ID POLG HPAVG STANDARD; PRT; 808 AA.
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
OS Hepatitis A virus (strain GA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2,

```

```

VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M66695; AAA45477.1; -
DR InterPro; IPR008975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN <1 2 COAT PROTEIN VP4 (PIA).
FT CHAIN 3 223 COAT PROTEIN VP2 (PIB).
FT CHAIN 224 470 COAT PROTEIN VP3 (PIC).
FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
FT CHAIN 771 >808 CORE PROTEIN P2A.
FT NON_TER 808 808
FT SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;
SQ
Query Match 100.0%; Score 102; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGKQRLKYAQEE 20
Db 778 PYKELRLEVGKQRLKYAQEE 797

RESULT 3
POLG HPAVC
ID POLG HPAVC STANDARD; PRT; 852 AA.
AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V., Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -!- SUBUNIT: The virus capsid is composed of one copy each of proteins VP1, VP2, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M10033; AAA45470.1; -
DR PIR; A03904; GNNYHA.
DR InterPro; IPR008975; Viral_cap_coat.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).

```


FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 >852 CORE PROTEIN P2A.
 FT NON_TER 852
 SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 852;
 Best Local Similarity 100.0%; Pred. No. 2.7e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 4
 POLG_HPAV2 STANDARD; PRT; 2226 AA.
 AC P26580;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE P3D (EC 2.7.7.48).
 OS Hepatitis A virus (strain 24a).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12094;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RA "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59810; AAA45468.1; --
 CC MEROPS; C03.005; --
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P3D.
 CC InterPro; IPR008975; Viral_cap_PSVir.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 KW CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 245 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 DB 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 5
 POLG_HPAV4 STANDARD; PRT; 2226 AA.
 AC P26581;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE P3D (EC 2.7.7.48).
 OS Hepatitis A virus (strain 43c).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12095;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Cromeans T., Jansen R.W.;
 RA "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59809; AAA45469.1; --
 CC MEROPS; C03.005; --
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P3D.
 CC InterPro; IPR008975; Viral_cap_PSVir.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 KW CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 245 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA9B09BF75 CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PYKELRLEVQKRLKYAQEE 20
 Db 799 PYKELRLEVQKRLKYAQEE 818

RESULT 6
 POLG HPAV8 STANDARD; PRT; 2226 AA.
 ID POLG HPAV8 STANDARD; PRT; 2226 AA.
 AC P26582;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain 18f).
 OS Hepatitis A virus positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12096;
 RN [1]
 PP SEQUENCE FROM N.A.
 RX MEDLINE=91162758; PubMed=1705995;
 RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
 RA Croweans T., Jansen R.W.;
 RA "Antigenic and genetic variation in cytopathic hepatitis A virus
 RT variants arising during persistent infection: evidence for genetic
 RT recombination.";
 RL J. Virol. 65:2056-2065(1991).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M59808; AAA45467.1; -;
 CC PDB; 1O47; 15-MAY-00.
 CC MEQOPS; C03.005; -;
 CC InterPro; IPR004004; Calici pol hel.
 CC InterPro; IPR009003; Cys Ser trypsin.
 CC InterPro; IPR000605; RNA helicase.
 CC InterPro; IPR007095; RNA pol DS PS.
 CC InterPro; IPR001205; RNA pol P3D.
 CC InterPro; IPR007094; RNA pol psvir.
 CC InterPro; IPR008975; Viral cap coat.
 CC Pfam; PF00680; RNA dep RNA pol; 1.
 CC Pfam; PF00910; RNA helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 KW Polyprotein; Coat protein; Core protein; Transferrase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).

FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
 FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
 FT CHAIN 492 794 COAT PROTEIN VP1 (PID).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
 SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;
 Query Match 100.0%; Score 102; DB 1; Length 2226;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PYKELRLEVQKRLKYAQEE 20
 Db 799 PYKELRLEVQKRLKYAQEE 818

RESULT 7
 POLG HPAVH STANDARD; PRT; 2227 AA.
 ID POLG HPAVH STANDARD; PRT; 2227 AA.
 AC P08617; P06443; Q81082;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain HM-175).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12098;
 RN [1]
 PP SEQUENCE FROM N.A.
 RX STRAIN=Wild type;
 RX MEDLINE=87061253; PubMed=3023706;
 RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
 RA Baroudy B.M.;
 RA "Complete nucleotide sequence of wild-type hepatitis A virus:
 RT comparison with different strains of hepatitis A virus and other
 RT picornaviruses.";
 RL J. Virol. 61:50-59(1987).
 RN [2]
 PP SEQUENCE FROM N.A.
 RX STRAIN=Attenuated;
 RX MEDLINE=87175701; PubMed=3031686;
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
 RA Purcell R.H.;
 RA "Complete nucleotide sequence of an attenuated hepatitis A virus:
 RT comparison with wild-type virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 RN [3]
 PP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RX MEDLINE=85166289; PubMed=2984684;
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
 RA Purcell R.H., Feinstone S.M.;
 RA "Sequence analysis of hepatitis A virus cDNA coding for capsid
 RT proteins and RNA polymerase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC [RNA] (N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
 CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; M14114; AAA45475.1; -;
 CC EMBL; M14707; AAA45465.1; -;
 CC EMBL; M14707; AAA45466.1; ALT_INIT.
 CC EMBL; M16632; AAA45471.1; -;
 CC PIR; A03905; A03905.
 CC PIR; A25981; GNNYHM.
 CC PIR; A94149; GNNYMK.
 CC PDB; 1HAV; 23-DEC-96.
 CC MEROPS; C03.005; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P3Vir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1087
 FT CHAIN 1088 1422
 FT CHAIN 1423 1496
 FT CHAIN 1497 1519
 FT CHAIN 1520 1738
 FT CHAIN 1739 2227
 FT VARIANT 77 77
 FT VARIANT 764 764
 FT VARIANT 821 821
 FT VARIANT 1052 1052
 FT VARIANT 1062 1062
 FT VARIANT 1118 1118
 FT VARIANT 1151 1151
 FT VARIANT 1163 1163
 FT VARIANT 1277 1277
 FT VARIANT 1500 1500
 FT VARIANT 1805 1805
 FT VARIANT 1930 1930
 FT VARIANT 1930 1930
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7ABE740A6 CRC64;

Query Match 100.0%; Score 102; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 |||||
 Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 8

ID POLG_HPAVL STANDARD; PRT; 2227 AA.
 AC P06441;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [contains: Coat proteins VP1 to VP4; Core proteins
 DE P2A to P2C; Probable proteins P3A to P3C; RNA-directed RNA polymerase

DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain LA).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85190549; PubMed=2986127;
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
 RA Merryweather J., van Nest G., Dina D.;
 RL "Primary structure and gene organization of human hepatitis A virus.";
 RT Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC {RNA}(N).
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; K02990; AAA45472.1; -;
 CC PIR; A03903; GNNYHR.
 CC MEROPS; C03.005; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P3Vir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUS.
 CC Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 23
 FT CHAIN 24 245
 FT CHAIN 246 491
 FT CHAIN 492 836
 FT CHAIN 837 980
 FT CHAIN 981 1076
 FT CHAIN 1077 1422
 FT CHAIN 1423 1484
 FT CHAIN 1485 1507
 FT CHAIN 1508 1578
 FT CHAIN 1579 2227
 FT CHAIN 1679 2227
 SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 102; DB 1; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 7.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
 |||||
 Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 9

ID POLG_HPAVM STANDARD; PRT; 2227 AA.
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

01-JAN-1990 (Rel. 13, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].

OS Simian hepatitis A virus (strain AGM-27).

OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.

OC NCBI_TaxID=121102;

[1]

SEQUENCE FROM N.A.

RP MEDLINE=91311420; PubMed=1649901;

RR MEDLINE=91311420; PubMed=1649901;

RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,

RA Purcell R.H.;

RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome

RT structure and growth in cell culture with other HAV strains.";

RL J. Gen. Virol. 72:1677-1683(1991).

[2]

SEQUENCE OF 1750-2164 FROM N.A.

RP MEDLINE=89232168; PubMed=2541023;

RR Balayan M.S., Kusov Y.Y., Andjapardize A.G., Tsarev S.A.,

RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;

RT "Variations in genome fragments coding for RNA polymerase in human

RT and simian hepatitis A viruses.";

RL FEBS Lett. 247:425-428(1989).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA} (N).

CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00924; BAA00766.1; -.

EMBL; X15461; CAA33490.1; -.

PIR; A30470; GNNYSA.

MEPROS; C03.005; -.

InterPro; IPR004004; Calici_pol_hel.

InterPro; IPR009003; Cys_Ser_typsin.

InterPro; IPR000605; RNA_helicase.

InterPro; IPR007095; RNA_pol_DS_PS.

InterPro; IPR001205; RNA_pol_P3D.

InterPro; IPR007094; RNA_pol_psvir.

InterPro; IPR008975; Viral_cap_coat.

Pfam; PF00680; RNA_dep_RNA_pol_1.

Pfam; PF00910; RNA_helicase; 1.

PRINTS; PR00918; CALICIVIRUSNS.

PolyProtein; Coat protein; Core protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.

CHAIN 1 27

FT CHAIN 28 249

FT CHAIN 250 495

FT CHAIN 496 795

FT CHAIN 796 984

FT CHAIN 985 1091

FT CHAIN 1092 1426

FT CHAIN 1427 1498

FT CHAIN 1499 1521

FT CHAIN 1522 1741

FT CHAIN 1742 2230

SEQUENCE 2230 AA; 251296 MW; 8782330E324E1F19 CRC64;

Query Match 93.1%; Score 95; DB 1; Length 2230;

Best Local Similarity 90.0%; Pred. No. 9.3e-07;

Matches	18;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps	0;
QY	1	PKYELRLEVGKQRLKYAEE	20						
DB	803	PKYELRMEVGKQRLKYAEE	822						
RESULT 11									
POLG	HPA	STANDARD;	PRT;	839	AA.				
ID	FOLG	HPA	STANDARD;	PRT;	839	AA.			
AC	P31788;								
DT	01-JUL-1993	(Rel. 26, Created)							
DT	01-JUL-1993	(Rel. 26, Last sequence update)							
DT	16-OCT-2001	(Rel. 40, Last annotation update)							
DE	Genome polyprotein	[Contains: Coat proteins VP1 TO VP4; Core protein							
DE	P2A]	(Fragment).							
OS	Simian hepatitis A virus	(strain CY-145).							
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;								
OC	Hepatovirus								
OX	NCBI_TaxID=31707;								
RP	[1]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=91311421;	PubMed=1649902;							
RA	Mainan O.V., Margolis H.S., Balayan M., Brinton M.A.;								
RT	"Sequence analysis of a new hepatitis A virus naturally infecting								
RL	Y. Gen. Virol. 72:1685-1689(1991).								
CC	-1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,								
CC	each of which is composed of one copy each of proteins VP1, VP2,								
CC	VP3, and VP4.								
CC	-1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcement/								
CC	or send an email to license@isb-sib.ch).								
CC	EMBL; M59286; AAA45473.1; --								
DR	PIR; JQ1180; GNNYS2.								
DR	InterPro; IPR008975; Viral_cap_coat.								
KW	Polyprotein; Coat protein; Core protein.								
FT	CHAIN 1	COAT PROTEIN VP4 (PIA).							
FT	CHAIN 23	COAT PROTEIN VP2 (PIB).							
FT	CHAIN 24	COAT PROTEIN VP3 (PIC).							
FT	CHAIN 246	COAT PROTEIN VP1 (PID).							
FT	CHAIN 492	COAT PROTEIN P2A.							
FT	CHAIN ?	>839							
FT	NON TER	839							
FT	SEQUENCE	839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;							
QY	Match	92.2%;	Score	94;	DB	1;	Length	839;	
DB	Best Local Similarity	90.0%;	Pred. No.	4.9e-07;	Indels	0;	Gaps	0;	
DB	Marches	18;	Conservative	1;	Mismatches	1;	Indels	0;	Gaps

[illegible]

RESULT 13

```

ILVE_RICPR          STANDARD;          PRT;          290 AA.
AC  005970;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Probable branched-chain amino acid aminotransferase (EC 2.6.1.42)
DE  (BCAT).
GN  ILVE OR RP428.
OS  Rickettsia prowazekii.
OC  Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC  Rickettsiaceae; Rickettsiella; Rickettsia.
OX  NCBI_TaxID=782;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Madrid E;
RX  MEDLINE=97419517; PubMed=9274032;
RA  Andersson J.O., Andersson S.G.B.;
RT  "Genomic rearrangements during evolution of the obligate
RT  intracellular parasite Rickettsia prowazekii as inferred from an
RT  analysis of 52015 bp nucleotide sequence.";
RL  Microbiology 143:2783-2795 (1997).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Madrid E;
RX  MEDLINE=99039499; PubMed=9823893;
RA  Andersson S.G.B., Zomorodipour A., Andersson J.O.,
RA  Sierhartz-Ponten T., Almark U.C.M., Podowski R.M., Naeslund A.K.,
RA  Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT  "The genome sequence of Rickettsia prowazekii and the origin of
RT  mitochondria.";
RL  Nature 396:133-140 (1998).
CC  -!- FUNCTION: Acts on leucine, isoleucine and valine (By similarity).
CC  -!- CATALYTIC ACTIVITY: L-leucine + 2-oxoglutarate = 4-methyl-2-
CC  oxopentanoate + L-glutamate.
CC  -!- COFACTOR: Pyridoxal phosphate.
CC  -!- PATHWAY: Valine and isoleucine biosynthesis.
CC  -!- SIMILARITY: Belongs to class-IV of pyridoxal-phosphate-dependent
CC  aminotransferases.
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
CC  EMBL; Y11777; CAA72450.1; -.
CC  DR  EMBL; AJ235271; CAA14885.1; -.
CC  DR  PIR; C71701; C71701.
CC  DR  HSP; P00510; IA3G.
CC  DR  InterPro; IPR001544; Aminotrans_IV.
CC  DR  Pfam; PF01063; aminotran_4; 1.
CC  DR  ProDom; PD001961; Aminotran_4; 1.
CC  DR  PROSITE; PS00770; AA_TRANSF_4; 1.
CC  DR  TRANSFAM; Aminotransferase; Branched-chain amino acid biosynthesis;
CC  Pyridoxal phosphate; Complete proteome.
CC  KW  BINDING 155 155 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC  SQ  SEQUENCE 290 AA; 33068 MW; 98374E325350763D CRC64;
CC
Query Match 43.1%; Score 44; DB 1; Length 290;
Best Local Similarity 61.1%; Pred. No. 13;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 3 KSLCLEVKGKRLKVAQE 20
DB 225 KSLCLEVKGKRLKVAQE 242

```

RESULT 14

```

FOJO DROME          STANDARD;          PRT;          583 AA.
AC  P54360; Q24176;
DT  01-OCT-1996 (Rel. 34, Created)
DT  01-OCT-1996 (Rel. 34, Last sequence update)
DT  15-MAR-2004 (Rel. 43, Last annotation update)
DE  Four-jointed protein.
GN  FJ.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=eye imaginal disk;
RX  MEDLINE=96038089; PubMed=7555705;
RA  Villano J.L., Katz F.N.;
RT  "four-jointed is required for intermediate growth in the proximal-
RT  distal axis in Drosophila.";
RL  Development 121:2767-2777 (1995).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Canton-S;
RX  MEDLINE=96187865; PubMed=8606003;
RA  Brodsky M.H., Steller H.;
RT  Positional information along the dorsal-ventral axis of the
RT  Drosophila eye: Graded expression of the four-jointed gene.";
RL  Dev. Biol. 173:428-446 (1996).
CC  -!- FUNCTION: Required for intermediate growth in the proximal-distal
CC  axis. May be required for cell-cell signaling during disk
CC  development.
CC  -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC
CC  EMBL; U28837; AAA69524.1; -.
CC  DR  EMBL; U44904; BAB01809.1; -.
CC  DR  FlyBase; FBgn0000658; fj.
CC  DR  GO; GO:0005886; C:plasma membrane; IDA.
CC  DR  GO; GO:0007267; P:cell-cell signaling; IDA.
CC  DR  GO; GO:0045198; P:establishment of epithelial cell polarity; IMP.
CC  DR  GO; GO:0007446; P:imaginal disc growth; IMP.
CC  DR  GO; GO:0016348; P:leg joint morphogenesis (sensu Holometabola); IMP.
CC  DR  GO; GO:0007474; P:wing vein specification; IMP.
CC  DR  GO; GO:0007474; P:wing vein specification; IMP.
CC  KW  Transmembrane; Signal-anchor.
CC  FT  DOMAIN 1 78 CYTOPLASMIC (POTENTIAL).
CC  FT  TRANSMEM 79 99 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC  (POTENTIAL).
CC  FT  DOMAIN 100 583 EXTRACELLULAR (POTENTIAL).
CC  FT  CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).
CC  FT  CARBOHYD 379 379 N-LINKED (GLCNAC. .) (POTENTIAL).
CC  FT  CARBOHYD 491 491 N-LINKED (GLCNAC. .) (POTENTIAL).
CC  FT  CONFLICT 126 126 T -> S (IN REF. 2).
CC  FT  CONFLICT 193 193 M -> L (IN REF. 2).
CC  FT  CONFLICT 288 288 P -> R (IN REF. 2).
CC  FT  CONFLICT 330 330 R -> A (IN REF. 2).
CC  SQ  SEQUENCE 583 AA; 65504 MW; 82F1EA2A299DB284 CRC64;
CC
Query Match 43.1%; Score 44; DB 1; Length 583;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 2 YKELRLEVGKRLKVAQE 19
DB 2 YDIXRLEAGQOKLQQAQ 19

```

DR EMBL; AJ289241; CAC10067.1; --
DR EMBL; AJ289241; CAC10068.1; --
DR EMBL; AJ289243; CAC10070.1; --
DR EMBL; BC028751; AAH28751.1; --
DR HSSP; P04574; 1ALV.
DR MEROPS; C02.017; --
DR MGI; 1891369; Capn12.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001300; Shprot_acsite.
DR InterPro; IPR000169; Shprot_acsite.
DR Pfam; PF00648; Peptidase C2; 1.
DR PRINTS; PR00704; CALPAIN_III; 1.
DR SMART; SM00720; calpain_III; 1.
DR SMART; SM00230; Cyspc; 1.
DR PROSITE; PS00203; CALPAIN_CAT; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00640; THIOL_PROTEASE_ASN; FALSE_NEG.
DR PROSITE; PS00139; THIOL_PROTEASE_CVS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; FALSE_NEG.
DR Hydrolase; Thiol protease; Multigene family;
KW Alternative splicing.
FT DOMAIN 45 341 CALPAIN CATALYTIC.
FT DOMAIN 342 541 DOMAIN III.
FT DOMAIN 542 720 DOMAIN IV.
FT CA_BIND 634 645 EF-HAND (POTENTIAL).
FT ACT_SITE 105 105 BY SIMILARITY.
FT ACT_SITE 259 259 BY SIMILARITY.
FT ACT_SITE 283 283 BY SIMILARITY.
FT VARSPPLIC 341 448 SEVLGSPAGGSGWHIFQGRVGRGSGGSGOPSAENFWT
FT VARSPPLIC 449 720 NQFRLTLEPDEEDDDDEEGPVGWGGAAGARGPARGGRV
FT VARSPPLIC 460 462 LLD -> GDR (in isoform 3).
FT VARSPPLIC 463 720 /FTID-VSP 007809.
FT VARSPPLIC 460 502 /FTID-VSP 007810.
FT VARSPPLIC 503 720 /FTID-VSP 007811.
FT VARSPPLIC 720 AA; 80588 MW; 37C07BDF0145B531 CRC64;
Query Match 43.1%; Score 44; DB 1; Length 720;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 1 PYKELRLEVGKQRLKYAQEE 20
Db 545 PYKLELELAQLFELAGEE 564
RESULT 16
GCP6_HUMAN
ID GCP6_HUMAN STANDARD; PRT; 1819 AA.
AC Q9ERT7; Q9BY91; Q9UGX3; Q9UGX4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Gamma-tubulin complex component 6 (GCP-6).
GN TUBGCP6 OR GCP6 OR KIAA1669.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RESULT 15

CC CANC_MOUSE STANDARD; PRT; 720 AA.
CC AC Q9ER56; Q9ER53; Q9ER54; Q9ER55;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 10-OCT-2003 (Rel. 42, Last annotation update)
CC DE Calpain 12 (EC 3.4.22.-).
CC GN CAPN12.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC [1]
CC RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
CC RX MEDLINE=20422668; PubMed=10964513;
CC RA Dear T.N., Meier N.T., Hunn M., Boehm T.;
CC RT "Gene structure, chromosomal localization and expression pattern of
CC RL Capn12, a new member of the calpain large subunit gene family.";
CC [2]
CC RN SEQUENCE FROM N.A. (ISOFORM 2).
CC RP TISSUE=Mammary gland;
CC RX MEDLINE=22388257; PubMed=12477932;
CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
CC RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
CC RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
CC RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
CC RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
CC RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
CC RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Mullahy S.J.,
CC RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullahy S.J.,
CC RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
CC RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
CC RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
CC RA Fahy J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
CC RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
CC RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
CC RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
CC RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.B.,
CC RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
CC RT "Generation and initial analysis of more than 15,000 full-length
CC human and mouse cDNA sequences.";
CC RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC CC -!- FUNCTION: Calcium-regulated non-lysosomal thiol-protease (By
CC similarity).
CC CC -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity.
CC CC -!- ALTERNATIVE PRODUCTS:
CC CC Event=Alternative splicing; Named isoforms=4;
CC CC Name=1;
CC CC IsoId=Q9ER56-1; Sequence=Displayed;
CC CC Name=2;
CC CC IsoId=Q9ER56-2; Sequence=VSP_007807, VSP_007808;
CC CC Name=3; Synonyms=variant 3;
CC CC IsoId=Q9ER56-3; Sequence=VSP_007809, VSP_007810;
CC CC Name=4; Synonyms=variant 2;
CC CC IsoId=Q9ER56-4; Sequence=VSP_007811, VSP_007812;
CC CC -!- TISSUE SPECIFICITY: Expression localized to the cortex of the hair
CC follicle during the anagen phase of hair cycle.
CC CC -!- SIMILARITY: Belongs to peptidase family C2.
CC CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC CC -!- SIMILARITY: Contains 1 Calpain catalytic domain.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; AJ289241; CAC10066.1; --

RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RX MEDLINE=2151508; PubMed=11694571;
RA Murphy S.M., Preble A.M., Patel U.K., O'Connell K.L., Dias D.P.,
RA Moritz M., Agard D., Stults J.T., Stearns T.;
RT "GCP5 and GCP6: two new members of the human gamma-tubulin complex.";
RL Mol. Biol. Cell 12:3340-3352(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Alnsough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beagley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Evans K.L., Fey J.M., Fleming K., French L., Durbin R.M., Ellington A.G.,
RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sultan J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shinomiya A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chiswick S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Oersky P., Rohlfing T.,
RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L.M., Mardis E., Waterston R.,
RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanaki J.P.,
RA Peyrard K., Kedra D., Seroussi E., Fransson I., Tapia I., Bruder C.E.,
RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
RA Khan A.S., Lane L., Tilahun Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [3]
RP SEQUENCE OF 33-1819 FROM N.A. (ISOFORM 2).
RX TISSUE=Brain;
RA MEDLINE=21156230; PubMed=11258795;
RA Hirosewa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.;
RT "Identification of novel transcribed sequences on human chromosome 22
by expressed sequence tag mapping.";
RL DNA Res. 8:1-9(2001).
CC -!- FUNCTION: Gamma-tubulin complex is necessary for microtubule
nucleation at the centrosome.
CC -!- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,
GCP3, GCP4, GCP5 and GCP6.
CC -!- SUBCELLULAR LOCATION: Centrosome.
CC -!- ALTERNATIVE PRODUCTS:
Event=alternative splicing, Named isoforms=2;
CC Name=1;
CC isoId=Q96R77-1; Sequence=Displayed;
CC Name=2;
CC isoId=Q96R77-2; Sequence=VSP_001624;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the GCP family.

CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
in positions 1371 and 1758.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF272887; AAK92968.1; -;
DR EMBL; AL022326; CAB83046.1; ALT_SEQ.
DR EMBL; AL022328; CAB83047.1; ALT_SEQ.
DR EMBL; AB051456; BAB33339.1; ALT_FRAME.
DR Genbank; HGNC:18127; TUBGCP6.
DR GO; GO:0008274; C:gamma-tubulin ring complex; IDA.
DR GO; GO:0008017; F:microtubule binding; IDA.
DR GO; GO:0007020; P:microtubule nucleation; IDA.
DR InterPro; IPR007259; Spc97_Spc98.
DR Pfam; PF04130; Spc97_Spc98; 1.
KW Microtubule; Repeat; Alternative splicing.
FT DOMAIN 1027 1269 9 X 27 AA TANDEM REPEATS.
FT REPEAT 1027 1053 1.
FT REPEAT 1054 1080 2.
FT REPEAT 1081 1107 3.
FT REPEAT 1108 1134 4.
FT REPEAT 1135 1161 5.
FT REPEAT 1162 1188 6.
FT REPEAT 1189 1215 7.
FT REPEAT 1216 1242 8.
FT REPEAT 1243 1269 9.
FT VARSPLIC 1724 1757 Missing (in isoform 2).
FT CONFLICT 567 567 S -> L (IN REF. 3).
FT CONFLICT 1377 1377 A -> T (IN REF. 2 AND 3).
FT CONFLICT 1621 1621 L -> V (IN REF. 2).
SQ SEQUENCE 1819 AA; 200455 MW; 923576544D34594A CRC64;
Query Match 42.6%; Score 43.5; DB 1; Length 1819;
Best Local Similarity 55.6%; Pred.No.1e+02;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
QY 3 KELRLVGVKQRL-KYAOE 19
DB 661 KELRLVGVKQRL-KYAOE 678
RESULT 17
UPP_AQUAE
ID UPP_AQUAE STANDARD; PRT; 208 AA.
AC O67914;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uracil phosphoribosyltransferase (EC 2.4.2.9) (UMP pyrophosphorylase)
DE (UPRase).
OS UPP OR URAP OR AQ_2163.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Kaller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
CC -!- CATALYTIC ACTIVITY: UMP + diphosphate = uracil + 5-phospho-
alpha-D-ribose 1-diphosphate.

SQ SEQUENCE 290 AA; 33243 MW; 86112A254A505177 CRC64;
Query Match 41.2%; Score 42; DB 1; Length 290;
Best Local Similarity 55.6%; Pred. No. 27;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QV 3 KELLEVGKORLYKQAE 20
| : | | : | | | |
225 KDLGLEVKERLKEQIE 242
RESULT 22
PARC ECOLI
ID PARC ECOLI STANDARD; PRT; 752 AA.
AC P20082; 069154;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Topoisomerase IV subunit A (EC 5.99.1.1).
GN PARC OR B3019 OR Z4373 OR ECS3903 OR SF3063 OR S3267.
OS Escherichia coli.
OS Escherichia coli O157:H7, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=9100424; PubMed=2170028;
RA Kato J.-I., Nishimura Y., Imamura R., Niki H., Hiraga S., Suzuki H.;
RT "New topoisomerase essential for chromosome segregation in E. coli";
RL Cell 63:393-404(1990).
RN [2]
RP ERRATUM.
RC SPECIES=E.coli; STRAIN=K12;
RA Kato J.-I., Nishimura Y., Imamura R., Niki H., Hiraga S., Suzuki H.;
RL Cell 65:1289-1290(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dinalanta E.T., Potamocous K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
RL Nature 409:529-533(2001).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Horda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [6]
RP SEQUENCE OF 724-752 FROM N.A.

RC SPECIES=E.coli;
RX MEDLINE=92212294; PubMed=1557036;
RA Coleman J.;
RT "Characterization of the Escherichia coli gene for
1-acyl-sn-glycerol-3-phosphate acyltransferase (plac).";
RL Mol. Gen. Genet. 232:295-303(1992).
RN [7]
RP REVISIONS, AND CHARACTERIZATION.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=94043292; PubMed=8227000;
RA Peng H., Marians K.J.;
RT "Escherichia coli topoisomerase IV. Purification, characterization,
subunit structure, and subunit interactions.";
RL J. Biol. Chem. 268:24481-24481(1993).
RN [8]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=2272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding X., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
RN [9]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
RX MEDLINE=22590274; PubMed=12704152;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RT "Complete genome sequence and comparative genomics of Shigella
flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
RN [10]
RP SEQUENCE OF 59-137 FROM N.A.
RC SPECIES=S.flexneri;
RX MEDLINE=99056880; PubMed=9867794;
RA Chu Y.W., Houang E.T.S., Cheng A.F.B.;
RT "Novel combination of mutations in the DNA gyrase and topoisomerase
IV genes in laboratory-grown fluoroquinolone-resistant Shigella
flexneri mutants.";
RL Antimicrob. Agents Chemother. 42:3051-3052(1998).
RN [11]
RP FUNCTION: Topoisomerase IV is essential for chromosome
segregation. It has relaxation of supercoiled DNA activity.
CC Performs the decatenation events required during the replication
of a circular DNA molecule.
CC -1- SUBUNIT: Composed of two subunits: parC and parE.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: STRONG, WITH THE A SUBUNIT OF GYRASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; M58408; AAA24297.1; ALT_INIT.
CC EMBL; M63491; AAA24396.1;
CC EMBL; U28377; AAG59187.1;
CC EMBL; AB000384; AAC76055.1;
CC EMBL; AB005531; AAG58155.1;
CC EMBL; AF002563; BAB37326.1;
CC EMBL; L22025; AAC36840.1;
CC EMBL; AB015318; AAN44541.1;
CC EMBL; AB016988; AAP1853.1;
CC EMBL; AF065132; AAC17115.1;
CC PIR; A65089; A65089.

DR PIR; G85961; G85961.
DR PIR; G91116; G91116.
DR HSSP; P09097; IAB4.
DR EcoGene; EG10686; parC.
DR InterPro; IPR005691; DNA_gyraseA_C.
DR InterPro; IPR002205; DNA_topoisomIV.
DR InterPro; IPR005742; TopoIV_A_Gneg.
DR Pfam; PF03989; DNA_gyraseA_C; 2.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
DR SMART; SM00434; TOP4C; 1.
DR TIGRFAMs; TIGR01062; parC_Gneg; 1.
KW Topoisomerase; Isomerase; DNA-binding; Complete proteome.
FT ACT_SITE 120 120 DNA CLEAVAGE (BY SIMILARITY).
FT CONFLICT 85 85 A -> P (IN REF. 10).
SQ SEQUENCE 752 AA; 83831 MW; OD4907B96CEE7086 CRC64;

Query Match 41.2%; Score 42; DB 1; Length 752;
Best Local Similarity 40.0%; Pred. No. 72;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 PYKELRELVGKQRLKYAQQE 20
DB 694 POSTLTHVGRKIKLRPEE 713

RESULT 23
PARC_SALTY STANDARD; PRT; 752 AA.
AC P26973;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Topoisomerase IV subunit A (EC 5.99.1.-).
GN PARC OR STM174
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=92089025; PubMed=1751451;
RA Luttinger A.L., Springer A.L., Schmid M.B.;
RA "A cluster of genes that affects nucleoid segregation in Salmonella
typhimurium";
RL New Biol. 3:687-697(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grawal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: Topoisomerase IV is essential for chromosome
segregation. It has relaxation of supercoiled DNA activity.
CC Performs the decatenation events required during the replication
of a circular DNA molecule.
CC -1- SUBUNIT: Composed of two subunits: parC and parE.
CC -1- SUBCELLULAR LOCATION: Membrane-associated.
CC -1- SIMILARITY: STRONG, WITH THE A SUBUNIT OF GYRASE.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch)

or send an email to license@isb-sib.ch).

CC EMBL; M68936; AAA27180.1; -.
CC EMBL; AB008846; AAL22048.1; -.
CC PIR; A45582; A45582.
CC HSSP; P09097; IAB4.
CC StyGene; SG10276; parC.
CC InterPro; IPR006691; DNA_gyraseA_C.
CC InterPro; IPR002205; DNA_topoisomIV.
CC InterPro; IPR005742; TopoIV_A_Gneg.
CC Pfam; PF03989; DNA_gyraseA_C; 2.
CC Pfam; PF00521; DNA_topoisomIV; 1.
CC ProDom; PD000742; DNA_topoisomIV; 1.
CC SMART; SM00434; TOP4C; 1.
CC TIGRFAMs; TIGR01062; parC_Gneg; 1.
KW Topoisomerase; Isomerase; DNA-binding; Complete proteome.
FT ACT_SITE 120 120 DNA CLEAVAGE (BY SIMILARITY).
FT CONFLICT 241 242 MR -> IG (IN REF. 1).
SQ SEQUENCE 752 AA; 84037 MW; F34FD7FFD20D6760 CRC64;

Query Match 41.2%; Score 42; DB 1; Length 752;
Best Local Similarity 40.0%; Pred. No. 72;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 PYKELRELVGKQRLKYAQQE 20
DB 694 POSTLTHVGRKIKLRPEE 713

RESULT 24
SYV_CHLTR STANDARD; PRT; 939 AA.
AC O84304;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Valyl-tRNA synthetase (EC 6.1.1.9) (Valine-tRNA ligase) (ValRS).
GN VALS OR CT302.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=613;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/Cx;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RA "Genome sequence of an obligate intracellular pathogen of humans:
Chlamydia trachomatis";
RL Science 282:754-759(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
+ L-valyl-tRNA(Val).
CC -1- SUBUNIT: Monomer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch)

CC EMBL; AE001302; AAC67895.1; -.
CC PIR; H71532; H71532.
CC HSSP; P96142; 1GAX.
CC InterPro; IPR002300; tRNA-synt_1a.
CC InterPro; IPR001412; tRNA-synt_1.
CC InterPro; IPR002303; tRNA-synt_val.
CC InterPro; IPR009008; VALRS_ILERS_edit.
CC Pfam; PF00133; tRNA-synt_1; 1.

DR PRINTS; PR00986; TRNASYNTHVAL.
DR TIGRFAMs; TIGR00422; VALS; 1.
DR PROSITE; PS00178; AA_TRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 47 57 "HIGH" REGION.
FT SITE 563 567 "XMSKS" REGION.
FT BINDING 566 566 ATP (BY SIMILARITY).
SQ SEQUENCE 939 AA; 107036 MW; CEB8449DC7BB9066 CRC64;
Query Match 41.2%; Score 42; DB 1; Length 939;
Best Local Similarity 38.9%; Pred. No. 90;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 3 KRLALEVGKQRLKYAQEE 20
DQ 97 RHLKASLGKORTDPSREE 114
RESULT 25
GBG5_HUMAN
ID GBG5_HUMAN STANDARD; PRT; 68 AA.
AC P30670; Q61015;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Guanine nucleotide-binding protein G(I)/G(S)/G(O) gamma-5 subunit.
GN GNG5 OR GNGT5.
OS Homo sapiens (Human),
OS Mus musculus (Mouse),
OS Rattus norvegicus (Rat), and
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606; 10090, 10116, 9913;
[1]
RN SEQUENCE FROM N.A.
RP SPECIES=Human;
RX MEDLINE=93003227; PubMed=9790912;
RA Liu B., Aronson N.N. Jr.;
RL "Structure of human G protein Ggamma5 gene GNG5.";
RN Biochem. Biophys. Res. Commun. 251:88-94(1998).
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=Human;
RX MEDLINE=9318631; PubMed=9553160;
RA Mao M., Fu G., Wu J.-S., Zhang Q.-H., Zhou J., Kan L.-X., Huang Q.-H.,
RA He K.-L., Gu B.-W., Han Z.-G., Shen Y., Gu J., Yu Y.-P., Xu S.-H.,
RA Wang Y.-X., Chen S.-J., Chen Z.;
RT "Identification of genes expressed in human CD34(+) hematopoietic
RT stem/progenitor cells by expressed sequence tags and efficient full-
RT length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:8175-8180(1998).
[3]
RN SEQUENCE FROM N.A.
RP SPECIES=Human;
RX MEDLINE=2238625; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Narusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faisey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RN SEQUENCE FROM N.A.
RP SPECIES=Bovine; TISSUE=Spleen;
RX MEDLINE=92195304; PubMed=1549114;
RA Fisher K.J., Aronson N.N. Jr.;
RT "Characterization of the cDNA and genomic sequence of a G protein
RT gamma subunit (gamma 5).";
RL Mol. Cell. Biol. 12:1585-1591(1992).
[6]
RN SEQUENCE.
RP SPECIES=Bovine; TISSUE=Spleen;
RX MEDLINE=93356792; PubMed=8352779;
RA Morishita R., Masuda K., Niwa M., Kato K., Asano T.;
RT "Identification of three forms of the gamma subunit of G proteins
RT isolated from bovine spleen.";
RL Biochem. Biophys. Res. Commun. 194:1221-1227(1993).
[7]
RN SEQUENCE OF 8-53 FROM N.A.
RP SPECIES=Mouse; STRAIN=CF-1 / Harlan;
RX MEDLINE=97011591; PubMed=8858601;
RA Williams C.J., Schultz R.M., Kopf G.S.;
RT "G protein gene expression during mouse oocyte growth and maturation,
RT and preimplantation embryo development.";
RL Mol. Reprod. Dev. 44:315-323(1996).
CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC involved as a modulator or transducer in various transmembrane
CC signaling systems. The beta and gamma chains are required for the
CC G-protein activity, for replacement of GDP by GTP, and for G protein-
CC effector interaction.
CC -1- SUBUNIT: G proteins are composed of 3 units, alpha, beta and
CC gamma.
CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues.
CC -1- SIMILARITY: Belongs to the G protein gamma family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF085709; AAC72203.1; -
CC EMBL; AF085708; AAC72203.1; JOINED.
CC EMBL; AF038955; AAC39869.1; -
CC EMBL; AF493873; AAM12587.1; -
CC EMBL; BC003563; AAR03563.1; -
CC EMBL; M95779; AAA30535.1; -
CC EMBL; M95780; AAA41188.1; -
CC EMBL; U38498; AAB01729.1; -
CC PIR; B42243; B42243.
CC Genew; HGNC:4408; GNG5.
CC MIM; 600974; -
CC MGD; MGI:109164; Gng5.
CC InterPro; IPR001770; G-gamma.
CC Pfam; PF00631; G-gamma; 1.
CC PRINTS; PR00321; GPROTEIN.
CC ProDom; PD003783; G-gamma; 1.
CC SMART; SM00224; GGL; 1.
CC SWART; PS00058; G-PROTEIN GAMMA; 1.
CC PROSITE; PS00058; G-PROTEIN GAMMA; 1.
KW Transducer; Prenylation; Lipoprotein; Multigene family.
FT Lipid 65 65 S-geranylgeranyl cysteine

```

FT      (By similarity).
SQ PROPEP      66      68      REMOVED IN MATURE FORM (BY SIMILARITY).
SQ SEQUENCE 68 AA; 7318 MW; 9AF7A16558683602 CRC64;

Query Match      40.2%; Score 41; DB 1; Length 68;
Best Local Similarity 50.08; Pred. No. 8.8;
Matches      Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      3 KKLRLVGVKQRLKYAQ 18
Db      :|||:|:|:|
      15 QQLRLVGLNVRKVQSQ 30

RESULT 26
GBGA HUMAN STANDARD; PRT; 68 AA.
ID AC
P50151.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Guanine nucleotide-binding protein G(1)/G(s)/G(o) gamma-10 subunit.
DE GNG10 OR NGNTG10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A., AND ISOPRENOID.
RP MEDLINE=95394940; PubMed=7665596;
RX Ray K., Kunsch C., Bonner L.M., Robishaw J.D.;
RT "Isolation of cDNA clones encoding eight different human G protein
RT gamma subunits, including three novel forms designated the gamma 4,
RT gamma 10, and gamma 11 subunits.";
RL J. Biol. Chem. 270:21765-21771(1995).
RN [2]
SEQUENCE FROM N.A.
RP Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
RA "cDNA clones of human proteins involved in signal transduction
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RP Laird G.;
RA Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RP TISSUE=Brain, and Lung;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulläh S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Munz D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchwood J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.T., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
FT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- FUNCTION: Guanine nucleotide-binding proteins (G proteins) are
CC involved as a modulator or transducer in various transmembrane
CC signaling systems. The beta and gamma chains are required for the
CC GTPase activity, for replacement of GDP by GTP, and for G protein-
CC effector interaction. Interacts with beta-1 and beta-2, but not

```

RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Distachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: CAN CATALYZE THE PRODUCTION OF LTC4 FROM LTA4 AND
CC REDUCED GLUTATHIONE. CAN CATALYZE THE CONJUGATION OF 1-CHLORO-2,4-
CC DINITROBENZENE WITH REDUCED GLUTATHIONE.
CC -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC -!- SUBUNIT: Homotrimer.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Liver, spleen, skeletal muscle, heart,
CC adrenals, pancreas, prostate, testis, fetal liver, and fetal
CC spleen. Very low expression in lung, brain, placenta and bone
CC marrow.
CC -!- SIMILARITY: Belongs to the MAPEG family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U77604; AAC15768.1; --
CC EMBL: BC025416; AAH25416.1; --
CC Genbank: HGNC:7063; MGST2.
CC MIM: 601733; --
CC GO: GO:0005624; C:membrane fraction; TAS.
CC GO: GO:0005792; C:microsome; TAS.
CC GO: GO:0004364; F:glutathione transferase activity; TAS.
CC GO: GO:0006960; P:antimicrobial humoral response (sensu Inver. .); TAS.
CC GO: GO:0007267; P:cell-cell signaling; TAS.
CC GO: GO:0006691; P:leukotriene metabolism; TAS.
CC GO: GO:0007185; P:signal transduction; TAS.
CC InterPro: IPR001446; S:lipoxynaseAP.
CC InterPro: IPR001129; MAPEG.
CC Pfam: PF01124; MAPEG; 1.
CC PRINTS: PR00488; 5LPOXGNASEAP.
CC PROSITE: PS01297; FLAP_GST2 LTC4s; 1.
KW Transferase; Transmembrane; Microsome; Leukotriene biosynthesis.
FT TRANSMEM 6 26 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 111 131 POTENTIAL.
SQ SEQUENCE 147 AA; 16620 MW; DDE89B46885D16EF CRC64;

Query Match 40.2%; Score 41; DB 1; Length 147;
Best Local Similarity 80.0%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 LEVGVKRLKY 16
:|||||
DB 24 LQVGKRLKY 33
:|||||

RESULT 28
YZ34_AQAAE STANDARD; PRT; 318 AA.
ID YZ34_AQAAE

AC 066423;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AA34.
GN AA34.
OS Aquifex aeolicus.
OC Plasmid ecoli.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5.
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.,
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus".
RT Nature 392:353-358 (1998).
CC -!- SIMILARITY: STRONG, TO A.EOLICUS AA07 AND AA11.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB000667; AAC07975.1; --
CC Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 318 AA; 37360 MW; EDB125C719DE49A3 CRC64;

Query Match 40.2%; Score 41; DB 1; Length 318;
Best Local Similarity 47.4%; Pred. No. 43;
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 YKELRLEVGKRLKYAQS 20
:|||||
DB 55 FFEETALWKQKYE 73
:|||||

RESULT 29
ATPD_MYCTU STANDARD; PRT; 446 AA.
ID ATPD_MYCTU
AC Q10594;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP synthase delta chain (EC 3.6.3.14).
GN ATPH OR RV1307 OR MT1347 OR MTC1373.27 OR MB1339.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773; 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broech R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala P.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence".
RT Nature 393:537-544 (1998).
RN [2]


```
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Emolaeva M., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W., Jacobs W.R. Jr., Venter J.C., Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RL laboratory strains";
RL J. Bacteriol. 184:5479-5490(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.bovis; STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972;
RA Garnier T., Eighmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duchoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
CC -!- FUNCTION: This protein seems to be part of the stalk that links
CC CF(0) to CF(1). It either transmits conformational changes from
CC CF(0) into CF(1) or is implicated in proton conduction.
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(2), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: a, b and c (by similarity).
CC -!- SIMILARITY: Belongs to the ATPase delta chain family.
CC -!- SIMILARITY: THIS PROTEIN IS MUCH LONGER THAN THAT OF OTHER
CC BACTERIAL DELTA CHAINS, THE C-TERMINAL PART IS HOMOLOGOUS TO DELTA
CC CHAINS WHILE THE N-TERMINAL REGION IS SIMILAR TO B/B' SUBUNITS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z73419; CAA97740.1; -.
CC EMBL; AE007008; AAX45609.1; -.
CC EMBL; BX248338; CAD94200.1; -.
CC PIR; G70774; G70774.
CC TIGR; MT1347; -.
CC TubercuList; Rv1307; -.
CC InterPro; IPR002146; ATPsynt_B/B'.sub.
CC InterPro; IPR000711; ATPsynt_OSCP.
CC Pfam; PF00430; ATP-synt B; 1.
CC Pfam; PF00213; OSCP; 1.
CC PRINTS; PR00125; ATPASEDELTA.
CC TIGRFAMs; TIGR01145; ATP_synt_delta; 1.
CC PROSITE; PS00389; ATPASE_DELTA; FALSE NEG.
CC HydroLase; ATP synthesis; CF(1); Hydrogen ion transport;
CC Complete proteome.
CC DOMAIN 1 150 ATP SYNTHASE B/B' SUBUNIT LIKE.
CC DOMAIN 261 446 ATP SYNTHASE DELTA CHAIN LIKE.
CC CONFLICT 25 25 L -> V (IN REF. 2).
CC SEQUENCE 446 AA; 48805 MW; AF07F26E3B878315 CRC64;
SQ
Query Match 40.2%; Score 41; DB 1; Length 446;
Best Local Similarity 41.2%; Pred. No. 61;
Matches 7; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
QY 3 KRLRLEVGKQRLKYAQE 19
DB 115 RQLRLGLGHESVRQARE 131
```

RESULT 30

```
MOP HAEIN
ID -MOP HAEIN STANDARD; PRT; 69 AA.
AC P45183;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable molybdenum-pterin binding protein.
GN H11370.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RL Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: BINDS ONE MOLE OF MOLYBDENUM PER MOLE OF PROTEIN AND
CC CONTAINS A PTERIN (BY SIMILARITY).
CC -!- SIMILARITY: TO C.PASTEURIANUM MOP PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32817; AAC23017.1; -.
CC PIR; I64119; I64119.
CC TIGR; H11370; -.
CC InterPro; IPR004606; Mop.
CC InterPro; IPR008995; MOP like.
CC InterPro; IPR005116; TOBE.
CC Pfam; PF03459; TOBE; 1.
CC TIGRFAMs; TIGR00638; Mop; 1.
CC Molybdenum; Complete proteome.
CC SEQUENCE 69 AA; 7122 MW; 1D32CE1C4C36310E CRC64;
SQ
Query Match 39.2%; Score 40; DB 1; Length 69;
Best Local Similarity 80.0%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 KRLRLEVGKQ 12
DB 45 KRLNLEVGKE 54
```

Search completed: May 11, 2004, 13:38:00
Job time : 8.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:28:26 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-09-171-432A-40

Perfect scores: 102

Sequence: 1 PYKELRLEVGKQRLKYAEE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: SPTEMBL 25.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mmc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	102	100.0	55	12 Q99972	Q99972 hepatitis a
2	102	100.0	55	12 Q99974	Q99974 hepatitis a
3	102	100.0	55	12 Q99973	Q99973 hepatitis a
4	102	100.0	55	12 Q99978	Q99978 hepatitis a
5	102	100.0	56	12 Q8JYP4	Q8JYP4 hepatitis a
6	102	100.0	56	12 Q67822	Q67822 hepatitis a
7	102	100.0	56	12 Q39872	Q39872 hepatitis a
8	102	100.0	56	12 Q99971	Q99971 hepatitis a
9	102	100.0	56	12 Q99978	Q99978 hepatitis a
10	102	100.0	56	12 Q8JYP2	Q8JYP2 hepatitis a
11	102	100.0	56	12 Q8JYP1	Q8JYP1 hepatitis a
12	102	100.0	56	12 Q99971	Q99971 hepatitis a
13	102	100.0	56	12 Q39867	Q39867 hepatitis a
14	102	100.0	56	12 Q8JYK3	Q8JYK3 hepatitis a
15	102	100.0	56	12 Q918Q3	Q918Q3 hepatitis a
16	102	100.0	56	12 Q91PB3	Q91PB3 hepatitis a

90 102 100.0 56 12 Q8JYT9
91 102 100.0 56 12 Q9IP94
92 102 100.0 56 12 Q9DIY5
93 102 100.0 56 12 Q8JYU0
94 102 100.0 56 12 Q9IPB5
95 102 100.0 56 12 Q9IPB6
96 102 100.0 56 12 Q9IP99
97 102 100.0 56 12 Q9DIX7
98 102 100.0 56 12 Q9IPAS
99 102 100.0 56 12 Q9IPB3
100 102 100.0 56 12 Q9IPCO

ALIGNMENTS

RESULT 1
Q999T2 ID Q999T2 PRELIMINARY; PRT; 55 AA.
AC Q999T2;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=Uru3;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
co-circulation during epidemic outbreaks."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306386; CAC29235.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 102; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKORLYAQEE 20
DB 35 PYKELRLEVGKORLYAQEE 54

RESULT 2
Q999T4 ID Q999T4 PRELIMINARY; PRT; 55 AA.
AC Q999T4;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-3;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;

RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
co-circulation during epidemic outbreaks."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306384; CAC29233.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 102; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKORLYAQEE 20
DB 35 PYKELRLEVGKORLYAQEE 54

RESULT 3
Q999T3 ID Q999T3 PRELIMINARY; PRT; 55 AA.
AC Q999T3;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-3;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
co-circulation during epidemic outbreaks."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306385; CAC29234.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 55
FT NON_TER 55 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 102; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKORLYAQEE 20
DB 35 PYKELRLEVGKORLYAQEE 54

RESULT 4
Q999U8 ID Q999U8 PRELIMINARY; PRT; 55 AA.
AC Q999U8;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

```
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg-6;
RA Costa-Mattoli M., Ferre V., Monphocho S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis A Virus in South America reveals
RT heterogeneity and co-circulation during epidemic outbreaks.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306370; CAC29219.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1
FT NON_TER 55
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 102; DB 12; Length 55;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 35 PYKELRLEVGKQRLKYAEE 54

RESULT 5
Q8JYP4 PRELIMINARY; PRT; 56 AA.
AC Q8JYP4
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Manuvelo;
RA Theamboonlers A., Jantaradsamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
RT outbreak in the southern part of Thailand.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507065; AAM33425.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 6
Q67822 PRELIMINARY; PRT; 56 AA.
AC Q67822
```

```
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ISR-70; Shulman L.M., Manor J., Leitner L., Shehab S.,
RA Karetnyi Y.V.;
RT "Genetic classification of hepatitis A virus strains isolated in
RT Israel, based on their VP1/2A nucleotide sequence.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z77247; CAB01040.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1
FT CHAIN 1 >28 CAPSID PROTEIN VP1.
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.
FT NON_TER 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
Db 36 PYKELRLEVGKQRLKYAEE 55

RESULT 7
O39872 PRELIMINARY; PRT; 56 AA.
AC O39872
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VDM;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RT 1982-1996.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68697; AAB53593.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAEE 20
```

```
Db 36 PYKELRLEVQKRLKYAQEE 55
|||||
ID Q999T1 PRELIMINARY; PRT; 56 AA.
AC Q999T1, 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H22;
RA Costa-Mattioli M., Ferre V., Monpholo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks.";
RL Submitted (FEF-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306387; CAC29236.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6602 MW; CE935CE05D46AEF4 CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVQKRLKYAQEE 20
|||||
Db 35 PYKELRLEVQKRLKYAQEE 54
|||||
ID Q999T8 PRELIMINARY; PRT; 56 AA.
AC Q999T8, 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-4;
RA Costa-Mattioli M., Ferre V., Monpholo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks.";
RL Submitted (FEF-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306380; CAC29229.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVQKRLKYAQEE 20
|||||
Db 35 PYKELRLEVQKRLKYAQEE 54
|||||
ID Q999T8 PRELIMINARY; PRT; 56 AA.
AC Q999T8, 2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chile-4;
RA Costa-Mattioli M., Ferre V., Monpholo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks.";
RL Submitted (FEF-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306380; CAC29229.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;
```

```
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVQKRLKYAQEE 20
|||||
Db 36 PYKELRLEVQKRLKYAQEE 55
|||||
ID Q8JYP2 PRELIMINARY; PRT; 56 AA.
AC Q8JYP2, 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAVAO;
RA Theamboonlers A., Jantaradamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
RT outbreak in the southern part of Thailand.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507067; AAM33427.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVQKRLKYAQEE 20
|||||
Db 36 PYKELRLEVQKRLKYAQEE 55
|||||
ID Q8JYP1 PRELIMINARY; PRT; 56 AA.
AC Q8JYP1, 2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maya;
RA Theamboonlers A., Jantaradamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
RT outbreak in the southern part of Thailand.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507068; AAM33428.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
```

DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 12
Q999U1 PRELIMINARY; PRT; 56 AA.
AC Q999U1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg-23;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
RA Billaudel S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
RT co-circulation during epidemic outbreaks."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306377; CAC29226.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000896; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 13
Q39867 PRELIMINARY; PRT; 56 AA.
AC Q39867;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JVR;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus:
RT 1982-1996."

J. Med. Virol. 51:273-279(1997).
RL EMBL; U68692; AAB53588.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000896; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 14
Q8JYK3 PRELIMINARY; PRT; 56 AA.
ID Q8JYK3
AC Q8JYK3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Yeongor 15;
RA Theamboonlers A., Jantaradsamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
RT outbreak in the southern part of Thailand."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF509835; AAM34763.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000896; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PYKELRLEVGVKQRLKYAQEE 20
DB 36 PYKELRLEVGVKQRLKYAQEE 55
RESULT 15
Q918Q3 PRELIMINARY; PRT; 56 AA.
ID Q918Q3
AC Q918Q3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=RJ-005;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 0:0-0(2001).
DR EMBL: AF410391; AAL10177.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0431; PQ0431.
DR InterPro: IPR000886; ER target S.
DR PROSITE: PS00014; ER_TARGET; 1.
FT NON_TER 1 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 16
Q9IPB3 PRELIMINARY; PRT; 56 AA.
AC Q9IPB3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC STRAIN=A306;
RA Fujiwara K.;
RL "hepatitis A virus VP1/2A junction."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB046897; BAB08054.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0430; PQ0430.
DR PIR: PQ0431; PQ0431.
DR InterPro: IPR000886; ER target S.
DR PROSITE: PS00014; ER_TARGET; 1.
FT NON_TER 1 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 17
Q9IPB2 PRELIMINARY; PRT; 56 AA.
AC Q9IPB2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.

OX NCBI_TaxID=12092;
RN [1]
RC STRAIN=RJ-005;
RA de Paula V.S., Baptista M.L., Lampe E., Niel C., Gaspar A.M.C.;
RT "Characterization of hepatitis A virus isolates from subgenotypes IA
and IB in Rio de Janeiro, Brazil."
RL J. Med. Virol. 0:0-0(2001).
DR EMBL: AF410391; AAL10177.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0430; PQ0430.
DR PIR: PQ0431; PQ0431.
DR InterPro: IPR000886; ER target S.
DR PROSITE: PS00014; ER_TARGET; 1.
FT NON_TER 1 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 18
Q999U7 PRELIMINARY; PRT; 56 AA.
AC Q999U7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RC STRAIN=Uru7;
RA Costa-Mattioli M., Ferre V., Monphoelo S., Garcia L., Colina R.,
RA Billandael S., Vega I., Perez-Bercoff R., Cristina J.;
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and
co-circulation during epidemic outbreaks."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ306371; CAC29220.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0430; PQ0430.
DR PIR: PQ0431; PQ0431.
DR InterPro: IPR000886; ER target S.
DR PROSITE: PS00014; ER_TARGET; 1.
FT NON_TER 1 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
DB 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 19
Q9IP95 PRELIMINARY; PRT; 56 AA.
AC Q9IP95;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=A7;
RA Fujiwara K.;
RL "hepatitis A virus VP1/2A junction.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046915; BAB08072.1; -
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1-
FT NON_TER 1 56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 20
Q99HM1 PRELIMINARY; PRT; 56 AA.
AC Q99HM1;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=KU94-420;
RA Song J.W., Byun K.S., Kim J.H., Song K.J., Baek L.J., Park S.H.,
RA Kwon O.S.;
RT "Molecular epidemiology of hepatitis A virus in Korea.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF234877; AAK00782.1; -
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1-
FT NON_TER 1 56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 21
O39870 PRELIMINARY; PRT; 56 AA.
ID O39870

O39870;
AC 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=406808;
RX MEDLINE=97247817; PubMed=9093940;
RA Taylor M.B.;
RT "Molecular epidemiology of South African strains of hepatitis A virus.";
RL J. Med. Virol. 51:273-279(1997).
DR EMBL; U68695; AAB53591.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1-
FT NON_TER 1 56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 22
Q99HL8 PRELIMINARY; PRT; 56 AA.
AC Q99HL8;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1] _
RP SEQUENCE FROM N.A.
RC STRAIN=KU98-54;
RA Song J.W., Byun K.S., Kim J.H., Song K.J., Baek L.J., Park S.H.,
RA Kwon O.S.;
RT "Molecular epidemiology of hepatitis A virus in Korea.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF234880; AAK00785.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1-
FT NON_TER 1 56
FT NON_TER 56 56
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PYKELRLEVGKQRLKYAQEE 20
Db 36 PYKELRLEVGKQRLKYAQEE 55

RESULT 23
O39870 PRELIMINARY; PRT; 56 AA.
ID O39870

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 23

Q91PAL PRELIMINARY; PRT; 56 AA.

AC Q91PAL; 01-OCT-2000 (TREMELrel. 15, Created)

DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=AS8;

RA Fujiwara K.;

RT "Hepatitis A virus VP1/2A junction.";

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB046909; BAB08066.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.

FT NON_TER 1

FT NON_TER 56

SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 7.7e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 24

O39866 PRELIMINARY; PRT; 56 AA.

AC O39866; 01-JAN-1998 (TREMELrel. 05, Created)

DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=96001190;

RX MEDLINE=97247817; PubMed=9093940;

RA Taylor M.B.;

RT "Molecular epidemiology of South African strains of hepatitis A virus: 1982-1996.";

RL J. Med. Virol. 51:273-279(1997).

DR EMBL; U68690; AAB53586.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.

FT NON_TER 1

FT NON_TER 56

SQ SEQUENCE 56 AA; 6614 MW; 8438C51846AEF4A4 CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 7.7e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 25

Q8JYK2 PRELIMINARY; PRT; 56 AA.

AC Q8JYK2; 01-OCT-2002 (TREMELrel. 22, Created)

DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Yeengor 16;

RA Theamboonlers A.; Jantaradamee P.; Poovorawan Y.;

RT "Molecular characterization of Hepatitis A virus infection of an outbreak in the southern part of Thailand.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF509836; AAM34764.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.

FT NON_TER 1

FT NON_TER 56

SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 7.7e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 26

Q999U6 PRELIMINARY; PRT; 56 AA.

AC Q999U6; 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ARG-4;

RA Costa-Wattoli M.; Perre V.; Monphoelo S.; Garcia L.; Colina R.;

RA Billaudel S.; Vega I.; Perez-Bercoff R.; Cristina J.;

RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and co-circulation during epidemic outbreaks.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ306372; CAC29221.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 25

Q8JYK2 PRELIMINARY; PRT; 56 AA.

AC Q8JYK2; 01-OCT-2002 (TREMELrel. 22, Created)

DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Yeengor 16;

RA Theamboonlers A.; Jantaradamee P.; Poovorawan Y.;

RT "Molecular characterization of Hepatitis A virus infection of an outbreak in the southern part of Thailand.";

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF509836; AAM34764.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.

FT NON_TER 1

FT NON_TER 56

SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;

Best Local Similarity 100.0%; Pred. No. 7.7e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20

Db 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 26

Q999U6 PRELIMINARY; PRT; 56 AA.

AC Q999U6; 01-JUN-2001 (TREMELrel. 17, Created)

DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polypeptide (Fragment)

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OX NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ARG-4;

RA Costa-Wattoli M.; Perre V.; Monphoelo S.; Garcia L.; Colina R.;

RA Billaudel S.; Vega I.; Perez-Bercoff R.; Cristina J.;

RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and co-circulation during epidemic outbreaks.";

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ306372; CAC29221.1; -

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

DR InterPro; IPR000886; ER target_S.

DR PROSITE; PS00014; ER_TARGET; 1.


```
FT NON_TER 1 1
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;
Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
DB 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 27
Q8JYP7 PRELIMINARY; PRT; 56 AA.
AC Q8JYP7;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=stMA;
RA Theamboonlers A., Jantaradsamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
outbreak in the southern part of Thailand."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507062; AA033422.1; -
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 1
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
DB 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 28
Q8JYP6 PRELIMINARY; PRT; 56 AA.
AC Q8JYP6;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB;
RA Theamboonlers A., Jantaradsamee P., Poovorawan Y.;
RT "Molecular characterization of Hepatitis A virus infection of an
outbreak in the southern part of Thailand."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF507063; AA033423.1; -
DR PIR; PQ0427; PQ0427.
```

```
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 1
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
DB 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 29
Q91P92 PRELIMINARY; PRT; 56 AA.
AC Q91P92;
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A75;
RA Fujiwara K.;
RT "Hepatitis A virus VP1/2A junction."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB046918; BAB08075.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
DR InterPro; IPR000886; ER_target_S.
DR PROSITE; PS00014; ER_TARGET; 1.
FT NON_TER 1 1
SQ SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVKGKRLKYAQEE 20
DB 36 PYKELRLEVKGKRLKYAQEE 55

RESULT 30
Q9DIX6 PRELIMINARY; PRT; 56 AA.
AC Q9DIX6;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE VP1 capsid protein and P2A protease (Fragment)
GN VP1 AND P2A.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CU-263F;
RX MEDLINE=21259955; PubMed=11360240;
```

```

RA Diaz B.I., Sariol C.A., Normann A., Rodriguez L.A., Flehmig B.;
RT "Genetic relatedness of Cuban HAV wild-type isolates.";
RL J. Med. Virol. 64:96-103 (2001).
DR EMBL; AJ245531; CAC17884.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR InterPro; IPR000886; ER target S.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Protease.
FT NON_TER 1 56
FT NON_TER 56
SQ SEQUENCE 56 AA; 6628 MW; 465CF4B35C1EF4BC CRC64;

Query Match 100.0%; Score 102; DB 12; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYKELRLEVGVKQLKYAQEE 20
   |||||
Db 36 PYKELRLEVGVKQLKYAQEE 55
   |||||

Search completed: May 11, 2004, 13:40:23
Job time : 34 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:37:37 ; Search time 34.25 Seconds
(without alignments)
162.083 Million cell updates/sec

Title: US-09-171-432A-40

Perfect score: 102
Sequence: 1 PYKELRLEVGKQRLKYAEE 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 27756755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	102	100.0	20	10	US-09-171-432A-40
2	102	100.0	352	14	US-10-272-459-45
3	102	100.0	836	14	US-10-272-459-40
4	102	100.0	980	14	US-10-272-459-41
5	102	100.0	2227	9	US-09-929-955-12
6	102	100.0	2227	13	US-10-104-966-12
7	102	100.0	2227	13	US-10-135-988-2
8	102	100.0	2227	13	US-10-135-988-4
9	102	100.0	2227	13	US-10-135-988-6
10	67	65.7	20	10	US-09-171-432A-39
11	46.5	45.6	65	11	US-09-864-408A-140
12	45	44.1	20	10	US-09-171-432A-41
13	45	44.1	25	10	US-09-171-432A-47
14	45	44.1	214	15	US-10-289-762-108
15	44	43.1	211	14	US-10-238-075-1284

16	44	43.1	270	12	US-10-424-599-252284	Sequence 252284,
17	44	43.1	720	12	US-10-312-373-1	Sequence 1, Appl
18	44	43.1	720	15	US-10-390-585-16	Sequence 16, Appl
19	44	43.1	720	16	US-10-055-569A-35	Sequence 35, Appl
20	44	43.1	720	16	US-10-055-569A-36	Sequence 36, Appl
21	44	43.1	720	16	US-10-055-569A-37	Sequence 37, Appl
22	43.5	42.6	1273	12	US-10-363-616-325	Sequence 325, Appl
23	43	42.2	120	14	US-10-238-075-999	Sequence 999, Appl
24	43	42.2	136	12	US-10-424-599-285976	Sequence 285976,
25	43	42.2	178	14	US-10-091-007-134	Sequence 124, Appl
26	42	41.2	90	12	US-10-424-599-216526	Sequence 216526,
27	42	41.2	156	12	US-10-424-599-226551	Sequence 226551,
28	42	41.2	207	14	US-10-238-075-1123	Sequence 1123, Appl
29	42	41.2	443	12	US-10-424-599-212457	Sequence 212457,
30	42	41.2	736	12	US-10-282-122A-76680	Sequence 76680, A
31	42	41.2	752	12	US-10-282-122A-56689	Sequence 56689, A
32	42	41.2	1054	12	US-10-001-885-108	Sequence 108, Appl
33	42	41.2	1104	12	US-10-425-114-57274	Sequence 57274, A
34	42	41.2	1122	12	US-10-412-699B-1582	Sequence 1582, Appl
35	42	41.2	4131	15	US-10-369-493-5136	Sequence 5136, Appl
36	41.5	40.7	358	15	US-10-369-493-3258	Sequence 3258, Appl
37	41.5	40.7	733	12	US-10-424-599-243766	Sequence 243766,
38	41	40.2	34	12	US-10-424-599-210256	Sequence 210256,
39	41	40.2	46	9	US-09-864-761-34339	Sequence 34339, A
40	41	40.2	68	14	US-10-161-941-18	Sequence 18, Appl
41	41	40.2	68	14	US-10-161-941-19	Sequence 19, Appl
42	41	40.2	69	9	US-09-982-809-5	Sequence 5, Appl
43	41	40.2	69	9	US-09-982-809-7	Sequence 7, Appl
44	41	40.2	77	12	US-10-424-599-189611	Sequence 189611,
45	41	40.2	79	9	US-09-925-301-877	Sequence 877, Appl
46	41	40.2	147	12	US-10-294-934-488	Sequence 488, Appl
47	41	40.2	147	14	US-10-247-671-133	Sequence 133, Appl
48	41	40.2	169	12	US-10-424-599-270895	Sequence 270895,
49	41	40.2	187	9	US-09-925-299-881	Sequence 881, Appl
50	41	40.2	187	10	US-09-925-299-881	Sequence 881, Appl
51	41	40.2	187	15	US-10-264-043-2421	Sequence 2421, Appl
52	41	40.2	337	12	US-10-424-599-255698	Sequence 255698,
53	41	40.2	397	14	US-10-168-425-13	Sequence 13, Appl
54	41	40.2	446	12	US-10-282-122A-62827	Sequence 62827, A
55	41	40.2	446	12	US-10-282-122A-64532	Sequence 64532, A
56	41	40.2	666	16	US-10-389-566-589	Sequence 589, Appl
57	41	40.2	767	15	US-10-369-493-9794	Sequence 9794, Appl
58	41	40.2	859	12	US-10-282-122A-65982	Sequence 65982, A
59	41	40.2	1137	12	US-10-425-114-63876	Sequence 63876, A
60	40.5	39.7	215	12	US-10-424-599-149292	Sequence 149292,
61	40	39.2	24	9	US-09-789-404-15	Sequence 15, Appl
62	40	39.2	171	12	US-10-424-599-160390	Sequence 160390,
63	40	39.2	182	14	US-10-191-513A-20	Sequence 20, Appl
64	40	39.2	189	9	US-09-847-209-2	Sequence 2, Appl
65	40	39.2	189	9	US-09-847-209-4	Sequence 4, Appl
66	40	39.2	189	15	US-10-437-581-2	Sequence 2, Appl
67	40	39.2	219	14	US-10-191-513A-19	Sequence 4, Appl
68	40	39.2	219	9	US-09-925-297-661	Sequence 661, Appl
69	40	39.2	272	14	US-10-191-513A-14	Sequence 14, Appl
70	40	39.2	287	14	US-10-191-513A-15	Sequence 15, Appl
71	40	39.2	288	14	US-10-425-114-45339	Sequence 45339, A
72	40	39.2	289	12	US-10-425-114-45339	Sequence 40, Appl
73	40	39.2	347	14	US-10-191-513A-40	Sequence 39, Appl
74	40	39.2	360	14	US-10-191-513A-39	Sequence 5, Appl
75	40	39.2	392	14	US-10-181-071-5	Sequence 2, Appl
76	40	39.2	419	9	US-09-948-774-2	Sequence 4556, Appl
77	40	39.2	425	12	US-10-296-115-1340	Sequence 1240, Appl
78	40	39.2	444	12	US-10-231-956A-69	Sequence 69, Appl
79	40	39.2	444	14	US-10-262-617-3	Sequence 3, Appl
80	40	39.2	444	14	US-10-191-513A-12	Sequence 12, Appl
81	40	39.2	444	14	US-10-191-513A-42	Sequence 42, Appl
82	40	39.2	444	16	US-10-429-160-22	Sequence 22, Appl
83	40	39.2	444	15	US-10-369-493-23388	Sequence 22388, A
84	40	39.2	470	15	US-10-289-762-143	Sequence 143, Appl
85	40	39.2	498	15	US-10-335-977-8259	Sequence 8259, Appl
86	40	39.2	501	12	US-10-335-977-8360	Sequence 8260, Appl
87	40	39.2	505	12	US-10-282-122A-72476	Sequence 72476, A
88	40	39.2	578	12		

; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41

Query Match 100.0%; Score 102; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 9.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 5

US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 102; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 6

US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US2002055124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03

; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 102; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 7

US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 102; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLEVGKQRLKYAQEE 20
Db 799 PYKELRLEVGKQRLKYAQEE 818

RESULT 8

US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6

```

; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1315
;
US-09-171-432A-39

Query Match 65.7%; Score 67; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PYKELRLGVGKQR 13
DB 8 PYKELRLGVGKQR 20

RESULT 11
US-09-864-408A-140
; Sequence 140, Application US/09964408A
; Publication No. US2004000947A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US2004000947A1 Human Polynucleotides and Polypept
; FILE REFERENCE: 21402-012
; CURRENT APPLICATION NUMBER: US/09/864,408A
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/206,690
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 9068
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 140
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-864-408A-140

Query Match 45.6%; Score 46.5; DB 11; Length 65;
Best Local Similarity 55.6%; Pred. No. 4.9;
Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 1 PYKELRLGVGKQR1KYAQ 18
DB 34 PYKGLR-EISRNMLRYAQ 50

RESULT 12
US-09-171-432A-41
; Sequence 41, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khuyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the

```

; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label= YK-1317
; US-09-171-432A-41

Query Match 44.1%; Score 45; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 QRLKYAQEE 20
DB 1 QRLKYAQEE 9

RESULT 13
US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudiyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1665
; US-09-171-432A-47

Query Match 44.1%; Score 45; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 QRLKYAQEE 20
DB 1 QRLKYAQEE 9

RESULT 14
US-10-289-762-108
; Sequence 108, Application US/10289762
; Publication No. US20040006218A1
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/10/289,762
; CURRENT FILING DATE: 2003-03-27
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 108
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; US-10-289-762-108

Query Match 44.1%; Score 45; DB 15; Length 214;
Best Local Similarity 50.0%; Pred. No. 32;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 KELRLEVGKORLYAQEE 20
DB 138 KYMAEVSKERLKLSE 155

RESULT 15
US-10-238-075-1284
; Sequence 1284, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; TITLE OF INVENTION: E.coli, and biological uses of these polynucleotides and of their
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075

US-10-312-373-1

Query Match 43.1%; Score 44; DB 14; Length 211;
Best Local Similarity 42.9%; Pred. No. 45;
Matches 9; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 1 PYKELRLEVGKORLKYAEE 19
Db 135 PYQIAREVKVPRRLKYLE 155

RESULT 16

US-10-424-599-252284
; Sequence 252284, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424, 599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 252284
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_69840C.1.p1ep

US-10-424-599-252284

Query Match 43.1%; Score 44; DB 12; Length 270;
Best Local Similarity 52.9%; Pred. No. 60;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 YKELRLVGVKORLKYAQ 18
Db 22 YKYFLIRVGVKORLREAR 38

RESULT 17

US-10-312-373-1

; Sequence 1, Application US/10312373
; Publication No. US20040072174A1
; GENERAL INFORMATION:
; APPLICANT: Boehm, Thomas
; APPLICANT: Dear, Neil T
; TITLE OF INVENTION: CALPAIN PROTEASE 12
; FILE REFERENCE: RKG-001US
; CURRENT APPLICATION NUMBER: US/10/312,373
; CURRENT FILING DATE: 2002-12-23
; PRIOR APPLICATION NUMBER: PCT/EP01/07457
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: DE 10031932.7
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Mus sp.

US-10-312-373-1

Query Match 43.1%; Score 44; DB 12; Length 720;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKORLKYAQEE 20
Db 545 PYKPLELELAQLFLELAGEE 564

RESULT 18

US-10-390-585-16
; Sequence 16, Application US/10390585
; Publication No. US20040014093A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN
; FILE REFERENCE: D0219DP
; CURRENT APPLICATION NUMBER: US/10/390,585
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: U.S. 60/364,941
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: US-10-390-585-16

US-10-390-585-16

Query Match 43.1%; Score 44; DB 15; Length 720;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKORLKYAQEE 20
Db 545 PYKPLELELAQLFLELAGEE 564

RESULT 19

US-10-055-569A-35
; Sequence 35, Application US/10055569A
; Publication No. US20040024181A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli, Esha A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gilbert, Jennifer
; APPLICANT: Casman, Stacie
; APPLICANT: Blalock, Angela
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit
; APPLICANT: Malyanker, Uriel
; APPLICANT: Stone, David
; APPLICANT: Millet, Isabelle
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Anderson, David W
; TITLE OF INVENTION: No. US20040024181A1 Human Proteins, Polynucleotides Encoding Th
; FILE REFERENCE: 21402-191
; CURRENT APPLICATION NUMBER: US/10/055,569A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,642

; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,681
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/243,863
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-055-569A-35

Query Match 43.1%; Score 44; DB 16; Length 720;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PYKLELEVGKQRLKYAQQE 20
DB 545 PYKLELELAQLFLELAGEE 564

RESULT 20
US-10-055-569A-36
; Sequence 36, Application US/10055569A
; Publication No. US20040024181A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli, Esha A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gilbert, Jennifer
; APPLICANT: Casman, Stacie
; APPLICANT: Blalock, Angela
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit
; APPLICANT: Malyanker, Uriel
; APPLICANT: Stone, David
; APPLICANT: Millet, Isabelle
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Anderson, David W
; TITLE OF INVENTION: No. US20040024181A1 Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-191
; CURRENT APPLICATION NUMBER: US/10/055,569A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,642
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592

; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,681
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/243,863
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 36
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-055-569A-36

Query Match 43.1%; Score 44; DB 16; Length 720;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 PYKLELEVGKQRLKYAQQE 20
DB 545 PYKLELELAQLFLELAGEE 564

RESULT 21
US-10-055-569A-37
; Sequence 37, Application US/10055569A
; Publication No. US20040024181A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli, Esha A
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gilbert, Jennifer
; APPLICANT: Casman, Stacie
; APPLICANT: Blalock, Angela
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Shenoy, Suresh
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit
; APPLICANT: Malyanker, Uriel
; APPLICANT: Stone, David
; APPLICANT: Millet, Isabelle
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Anderson, David W
; TITLE OF INVENTION: No. US20040024181A1 Human Proteins, Polynucleotides Encoding Th
; TITLE OF INVENTION: Methods of Using the Same
; FILE REFERENCE: 21402-191
; CURRENT APPLICATION NUMBER: US/10/055,569A
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: 60/243,642
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,320
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,592
; PRIOR FILING DATE: 2000-10-26
; PRIOR APPLICATION NUMBER: 60/243,681
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/243,863

```
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-055-569A-37

Query Match      43.1%; Score 44; DB 16; Length 720;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy      1 PYKELRELVGKQRLKYAQEE 20
Db      545 PYKPLELELAQLELAGEE 564

RESULT 22
US-10-363-616-325
; Sequence 325, Application US/10363616
; Publication No. US20040044181A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-113 (793)
; CURRENT APPLICATION NUMBER: US/10/363,616
; CURRENT FILING DATE: 2003-03-03
; PRIOR APPLICATION NUMBER: 09/654,935
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 490
; SEQ ID NO 325
; LENGTH: 1273
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-363-616-325

Query Match      42.6%; Score 43.5; DB 12; Length 1273;
Best Local Similarity 55.6%; Pred. No. 4e+02;
Matches 10; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy      3 KELRELVGKQRL-KYAE 19
Db      109 KELRMEIAKQELIAHARE 126

RESULT 23
US-10-238-075-999
; Sequence 999, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolated
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 999

; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 720
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-055-569A-37

Query Match      42.2%; Score 43; DB 14; Length 120;
Best Local Similarity 47.6%; Pred. No. 35;
Matches 10; Conservative 5; Mismatches 4; Indels 2; Gaps 1;

Qy      1 PYKELRELV--GKQRLKYAQEE 19
Db      67 PYQVAREVKAPREKRLKYLE 87

RESULT 24
US-10-424-599-265976
; Sequence 265976, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265976
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_82197C.1.pep
; US-10-424-599-265976

Query Match      42.2%; Score 43; DB 12; Length 136;
Best Local Similarity 45.0%; Pred. No. 41;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy      1 PYKELRELVGKQRLKYAQEE 20
Db      75 PYKGLDGDGVDGLHYCPDE 94

RESULT 25
US-10-091-007-124
; Sequence 124, Application US/10091007
; Publication No. US20030170782A1
; GENERAL INFORMATION:
; APPLICANT: Microbial Technics limited
; APPLICANT: Le Page, Richard W F
; APPLICANT: Hamnify, Sean B
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21978WO
; CURRENT APPLICATION NUMBER: US/10/091,007
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: GB 9921125.2
; PRIOR FILING DATE: 1999-09-07
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 124
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Streptococcus agalactiae
; US-10-091-007-124

Query Match      42.2%; Score 43; DB 14; Length 178;
```

Best Local Similarity 58.3%; Pred. No. 55;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKQ 12
DB 117 PFKELRLSAGRE 128

RESULT 26
US-10-424-599-216526
; Sequence 216526, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216526
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37551C.1.pep
US-10-424-599-216526

Query Match 41.2%; Score 42; DB 12; Length 90;
Best Local Similarity 42.1%; Pred. No. 38;
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 PYKELRLVGVKQRLKYAQE 19
DB 48 PERSIGLPLKHLRYVQE 66

RESULT 27
US-10-424-599-226551
; Sequence 226551, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 226551
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(156)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_46604C.1.pep
US-10-424-599-226551

Query Match 41.2%; Score 42; DB 12; Length 156;
Best Local Similarity 52.9%; Pred. No. 69;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 YKELRLVGVKQRLKYAQ 18

Db 69 YGELRPPLGKHLKIVE 85

RESULT 28
US-10-238-075-1123
; Sequence 1123, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
; APPLICANT: I.N.S.E.R.M.
; TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolate
; TITLE OF INVENTION: E.coli, and biological uses of these polynucleotides and of their
; FILE REFERENCE: BLANDINE
; CURRENT APPLICATION NUMBER: US/10/238,075
; CURRENT FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 0003145
; PRIOR FILING DATE: 2000-03-10
; NUMBER OF SEQ ID NOS: 1576
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1123
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Escherichia coli
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33874C.1.pep
US-10-238-075-1123

Query Match 41.2%; Score 42; DB 14; Length 207;
Best Local Similarity 52.4%; Pred. No. 94;
Matches 11; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

QY 1 PYKELLEV--GKQRLKYAQE 19
DB 131 PYGLIAREVKAPQRLKYALE 151

RESULT 29
US-10-424-599-212457
; Sequence 212457, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212457
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33874C.1.pep
US-10-424-599-212457

Query Match 41.2%; Score 42; DB 12; Length 443;
Best Local Similarity 53.3%; Pred. No. 2.2e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 6 RLEVGVKQRLKYAQEE 20
DB 112 RLNVGKEQVQAQNE 126

RESULT 30
US-10-282-122A-76680
; Sequence 76680, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 76680
LENGTH: 736
TYPE: PRT
ORGANISM: Ureaplasma urealyticum
US-10-282-122A-76680
Query Match 41.2%; Score 42; DB 12; Length 736;
Best Local Similarity 44.4%; Pred. No. 3.8e+02;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 2 YKELELVGKQRLKYAQE 19
DB 396 YKAIYSEINKKNQYLQE 413
Search completed: May 11, 2004, 13:44:49
Job time : 35.25 secs

GenCore version 5.1.6 Copyright (c) 1993 - 2004 CompuGen Ltd.																	
OM protein - protein search, using sw model																	
Run on:	May 11, 2004, 13:23:11	Search time 47 Seconds (without alignments) 120.233 Million cell updates/sec															
Title:	US-09-171-432A-41																
Perfect score:	103																
Sequence:	1 QRLKYAQEELSNEVLPPPRK 20																
Scoring table:	BLOSUM62 Gapop 10.0 , Gapext 0.5																
Searched:	1586107 seqs, 282547505 residues																
Total number of hits satisfying chosen parameters:	1586107																
Minimum DB seq length:	0																
Maximum DB seq length:	2000000000																
Post-processing:	Minimum Match 0% Maximum Match 100% Listing first 100 summaries																
Database :	A_Geneseq_29Jan04.* 1: Geneseqp1980s.* 2: Geneseqp1990s.* 3: Geneseqp2000s.* 4: Geneseqp2001s.* 5: Geneseqp2002s.* 6: Geneseqp2003as.* 7: Geneseqp2003bs.* 8: Geneseqp2004s.*																
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.																	
SUMMARIES																	
Result No.	Score	Query Match %	Length	DB ID	Description												
1	103	100.0	20	2	AAW42924 Immunogen												
2	103	100.0	21	4	AAB69441 Synthetic												
3	103	100.0	25	2	AAW42930 Immunogen												
4	103	100.0	26	4	AAB69447 Synthetic												
5	103	100.0	366	1	AAP50230 Sequence												
6	103	100.0	854	1	AAP50287 Hepatitis												
7	103	100.0	993	1	AAP50116 Sequence												
8	103	100.0	993	1	AAP50231 Sequence												
9	103	100.0	1077	2	AAW95559 A partial												
10	103	100.0	1091	2	AAR32426 Translate												
11	103	100.0	2227	1	AAP60066 Sequence												
12	103	100.0	2227	2	AAR05697 Attenuate												
13	103	100.0	2227	2	AAW34074 Hepatitis												
14	103	100.0	2227	3	AAB18609 Amino aci												
15	103	100.0	2227	3	AAB18607 Hepatitis												
16	103	100.0	2227	5	AAB18609 Hepatitis												
17	103	100.0	2227	5	ABG31729 Attenuate												
18	103	100.0	2227	5	ABG31727 Wild-type												
19	103	100.0	2227	6	ABU08641 Attenuat												
20	103	100.0	2227	6	ABU08639 Wild type												
21	103	100.0	2227	7	ABW00350 Hepatitis												
22	98	95.1	2227	3	AAB18608 Amino aci												
23	98	95.1	2227	5	ABG31728 Hepatitis												
24	98	95.1	2227	6	ABU08640 Attenuat												
25	94	91.3	839	2	AAR15629 Capsid re												

26	50	48.5	150	4	ABB68507	Drosophil
27	48	46.6	444	2	AAW20624	H. pylori
28	48	46.6	1176	4	ABB59692	Drosophil
29	47	45.6	221	4	ABB68504	Drosophil
30	45	43.7	20	2	AAW42923	Immunogen
31	45	43.7	21	4	AAB69440	Synthetic
32	45	43.7	263	4	AAW18008	Peptide #
33	45	43.7	263	4	ABB37043	Peptide #
34	45	43.7	263	4	AAW30519	Peptide #
35	45	43.7	263	4	ABB31812	Peptide #
36	45	43.7	263	4	ABB22358	Protein #
37	45	43.7	263	4	AAW70185	Human bon
38	45	43.7	263	4	AAW57770	Human bra
39	45	43.7	263	4	ABG51885	Human liv
40	45	43.7	263	4	AAW05648	Human pep
41	45	43.7	263	5	ABG39819	Human pol
42	45	43.7	305	4	AAW41677	Human lun
43	45	43.7	1596	5	AAE20271	Human lun
44	44	42.7	144	4	ABG26758	Novel hum
45	44	42.7	169	7	ADC31218	Human nov
46	44	42.7	169	7	ADC32911	Human nov
47	44	42.7	334	4	ABG26759	Novel hum
48	44	42.7	368	4	ABG28311	Novel hum
49	44	42.7	420	4	ABB63362	Drosophil
50	44	42.7	450	4	ABG26763	Novel hum
51	44	42.7	531	5	ABJ11004	Yeast sel
52	44	42.7	553	6	ABP72626	Hyphartri
53	44	42.7	533	4	ABG28306	Novel hum
54	44	42.7	755	4	ABB59259	Drosophil
55	44	42.7	914	2	AAW72724	ORC1 subu
56	44	42.7	914	2	AAW22224	S. cerevi
57	44	42.7	914	6	ABR53642	Partial s
58	43.5	42.2	1091	1	AAW98500	Human Duc
59	43.5	42.2	3685	1	AAW90373	Sequence
60	43.5	42.2	3685	1	AAW90290	Human gen
61	43	41.7	205	4	AAE07109	Human gen
62	43	41.7	213	4	AAE07083	Human gen
63	43	41.7	213	5	ABG65103	Human alb
64	43	41.7	302	3	AAW06817	Arabidops
65	43	41.7	362	3	AAW06816	Arabidops
66	43	41.7	372	3	AAW06815	Arabidops
67	43	41.7	390	2	AAW05302	S. aureus
68	43	41.7	584	5	AAE23121	Influenza
69	43	41.7	585	2	AAW01671	Influenza
70	43	41.7	585	2	AAW75443	Influenza
71	43	41.7	585	4	AAE04953	Influenza
72	43	41.7	586	2	AAW01675	Influenza
73	43	41.7	586	2	AAW75447	Influenza
74	43	41.7	586	4	AAE04957	Influenza
75	43	41.7	589	2	AAW01672	Influenza
76	43	41.7	589	2	AAW75444	Influenza
77	43	41.7	589	4	AAE04954	Influenza
78	43	41.7	592	2	AAW01674	Influenza
79	43	41.7	592	2	AAW75446	Influenza
80	43	41.7	592	4	AAE04956	Influenza
81	43	41.7	869	4	AAU33734	Staphyloc
82	43	41.7	869	6	ABU16498	Protein e
83	43	41.7	869	6	ABM73223	Staphyloc
84	43	41.7	882	4	AAU36933	Staphyloc
85	43	41.7	916	4	ABB63615	Drosophil
86	43	41.7	1058	6	ADA54341	Human pro
87	43	41.7	1095	6	ABJ25381	Human pro
88	43	41.7	1277	6	ABJ25381	Aspergill
89	43	41.7	1322	4	ABG60275	Aspergill
90	43	41.7	1345	3	ABJ18284	Plasmodi
91	43	41.7	1605	4	ABF70375	Drosophil
92	43	41.7	2357	7	ADC31432	Human nov
93	43	41.7	2696	6	AAE33492	Human REM
94	43	41.7	2696	6	AAO27402	Human NSD
95	43	41.7	2696	6	AAO27404	Human NSD
96	42.5	41.3	86	3	AG03559	Human sec
97	42.5	41.3	241	7	ADB64347	Human pro
98	42.5	41.3	468	7	ADB65354	Human pro

99 42.5 41.3 530 3 AAB42529 Human ORF
100 42.5 41.3 530 5 ABP68940 Human pol

ALIGNMENTS

RESULT 1

AAW42924
ID AAW42924 standard; peptide; 20 AA.

XX AC AAW42924;

XX DT 28-APR-1998 (first entry)

XX DE Immunogenic Hepatitis A virus peptide YK-1317.

XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.

XX OS Synthetic.

XX OS Hepatitis A virus.

XX PN WO9740147-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006891.

XX PR 19-APR-1996; 96US-0015644P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 1997-535831/49.

XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
XX response to HAV in a mammal or to detect the presence of antibodies
XX against HAV in a mammal.

XX PS Claim 18; Page 112; 140pp; English.

XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to
XX immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
XX substantially similar to a portion of the amino acid sequence of the P2A
XX protein of HAV corresponding to amino acids 792-980. The present peptide
XX is derived from amino acids 810-829, and has a reactivity of 83.3% with
XX acute sera. Compositions containing the peptides can be used to induce an
XX immune response to HAV in a mammal. The peptides can also be used to
XX detect the presence of antibodies against HAV in mammalian serum. The
XX peptides can also be used to make an antibody against HAV by
XX administering the peptide to a mammal

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 103; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.9e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

DB 1 QRLKYAQEELSNEVLPPPRK 20

RESULT 2

AAB69441

ID AAB69441 standard; peptide; 21 AA.

XX AC AAB69441;

XX DT 20-APR-2001 (first entry)

XX PN WO9740147-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006891.

XX PR 19-APR-1996; 96US-0015644P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 1997-535831/49.

XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
XX response to HAV in a mammal or to detect the presence of antibodies
XX against HAV in a mammal.

XX PS Claim 18; Page 112; 140pp; English.

XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to
XX immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
XX substantially similar to a portion of the amino acid sequence of the P2A
XX protein of HAV corresponding to amino acids 792-980. The present peptide
XX is derived from amino acids 810-829, and has a reactivity of 83.3% with
XX acute sera. Compositions containing the peptides can be used to induce an
XX immune response to HAV in a mammal. The peptides can also be used to
XX detect the presence of antibodies against HAV in mammalian serum. The
XX peptides can also be used to make an antibody against HAV by
XX administering the peptide to a mammal

XX SQ Sequence 20 AA;

Query Match 100.0%; Score 103; DB 2; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.9e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

DB 1 QRLKYAQEELSNEVLPPPRK 20

RESULT 3

AAW42930

ID AAW42930 standard; peptide; 25 AA.

XX AC AAW42930;

XX DT 28-APR-1998 (first entry)

XX DE Immunogenic Hepatitis A virus peptide YK-1665.

XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.

XX OS Synthetic.

XX OS Hepatitis A virus.

XX PN WO9740147-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006891.

XX PR 19-APR-1996; 96US-0015644P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 2001-112681/12.

XX PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.

XX PS Claim 13; Page 94; 130pp; English.

XX CC The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IGM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or
XX electrophoresis. The peptides are used to induce an immune response to
XX HAV when administered to a human or animal. Glutamine at the carboxy end
XX of the peptides enhances the IGM antibody reactivity

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 103; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 9.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

DB 1 QRLKYAQEELSNEVLPPPRK 20

RESULT 3

AAW42930

ID AAW42930 standard; peptide; 25 AA.

XX AC AAW42930;

XX DT 28-APR-1998 (first entry)

XX DE Immunogenic Hepatitis A virus peptide YK-1665.

XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.

XX OS Synthetic.

XX OS Hepatitis A virus.

XX PN WO9740147-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006891.

XX PR 19-APR-1996; 96US-0015644P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 2001-112681/12.

XX PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX detecting anti-hepatitis A virus and as vaccines.

XX PS Claim 13; Page 94; 130pp; English.

XX CC The present sequence is one of a number of synthetic peptides which are
XX immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX comprise antigenic epitopes of the major structural capsid polypeptides
XX or non-structural polypeptides of HAV with one or more glutamine
XX molecules at the carboxy end of the peptide. The peptides are used to
XX detect the presence of antibodies against HAV in mammalian serum, to
XX detect the presence of HAV in a human or animal through the binding of
XX the peptide to an antibody, to detect acute phase infection by detecting
XX IGM antibodies in mammalian serum and detecting convalescence in a
XX mammal. The peptides are used to detect or quantify HAV antibodies in
XX samples in clinical or research-based assays using immunoblotting,
XX fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX tracking of radioactive or bioluminescent markers, chromatography or
XX electrophoresis. The peptides are used to induce an immune response to
XX HAV when administered to a human or animal. Glutamine at the carboxy end
XX of the peptides enhances the IGM antibody reactivity

XX SQ Sequence 21 AA;

Query Match 100.0%; Score 103; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 9.4e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

DB 1 QRLKYAQEELSNEVLPPPRK 20

PD 30-OCT-1997.
 XX
 PF 18-APR-1997; 97WO-US006891.
 XX
 PR 19-APR-1996; 96US-0015644P.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Fields HA, Khudyakov YE;
 XX
 DR WPI; 1997-535831/49.
 XX
 XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
 PT response to HAV in a mammal or to detect the presence of antibodies
 PT against HAV in a mammal.
 XX
 PS Claim 18; Page 112; 140pp; English.
 XX
 XX Peptides AAW42922-30 are immunogenic peptides corresponding to
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
 CC substantially similar to a portion of the amino acid sequence of the P2A
 CC protein of HAV corresponding to amino acids 792-980. Compositions
 CC containing the peptides can be used to induce an immune response to HAV
 CC in a mammal. The peptides can also be used to detect the presence of
 CC antibodies against HAV in mammalian serum. The peptides can also be used
 CC to make an antibody against HAV by administering the peptide to a mammal
 XX
 SQ Sequence 25 AA;
 Query Match 100.0%; Score 103; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 1.1e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 DB 1 QRLKYAQEELSNEVLPPPRK 20
 RESULT 4
 AAB69447
 ID AAB69447 standard; peptide; 26 AA.
 XX
 AC AAB69447;
 XX
 DT 20-APR-2001 (first entry)
 XX
 DE Synthetic HAV P2A peptide, SEQ ID NO: 47.
 XX
 KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
 KW antigen; major structural capsid polypeptide; HAV antibody detection.
 XX
 OS Hepatitis A virus.
 OS Synthetic.
 PN WO200105824-A2.
 XX
 XX 25-JAN-2001.
 PD
 PF 14-JUL-2000; 2000WO-US019267.
 XX
 PR 15-JUL-1999; 99US-0144412P.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA
 PI Fields HA, Khudyakov YE;
 XX
 XX WPI; 2001-112681/12.
 DR
 XX Synthetic peptides used as antigen sources for enzyme immunoassays
 PT detecting anti-hepatitis A virus and as vaccines.
 XX
 PS Claim 13; Page 98; 130pp; English.
 XX

CC The present sequence is one of a number of synthetic peptides which are
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
 CC comprise antigenic epitopes of the major structural capsid polypeptides
 CC or non-structural polypeptides of HAV with one or more glutamine
 CC molecules at the carboxy end of the peptide. The peptides are used to
 CC detect the presence of antibodies against HAV in mammalian serum, to
 CC detect the presence of HAV in a human or animal through the binding of
 CC the peptide to an antibody, to detect acute phase infection by detecting
 CC IGM antibodies in mammalian serum and detecting convalescence in a
 CC mammal. The peptides are used to detect or quantify HAV antibodies in
 CC samples in clinical or research-based assays using immunoblotting,
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
 CC tracking of radioactive or bioluminescent markers, chromatography or
 CC electrophoresis. The peptides are used to induce an immune response to
 CC HAV when administered to a human or animal. Glutamine at the carboxy end
 CC of the peptides enhances the IGM antibody reactivity
 XX
 SQ Sequence 26 AA;
 Query Match 100.0%; Score 103; DB 4; Length 26;
 Best Local Similarity 100.0%; Pred. No. 1.2e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 DB 1 QRLKYAQEELSNEVLPPPRK 20
 RESULT 5
 AAP50230
 ID AAP50230 standard; protein; 366 AA.
 XX
 AC AAP50230;
 XX
 DT 28-NOV-1991 (first entry)
 XX
 DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
 XX
 KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
 KW diagnostic assay.
 XX
 OS Hepatitis A virus.
 XX
 PN EP138704-A.
 XX
 PD 24-APR-1985.
 XX
 PF 09-OCT-1984; 84EP-00402025.
 XX
 PR 14-OCT-1983; 83US-00541836.
 PR 02-MAR-1984; 84US-00585942.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 XX Hughes JV, Scolnick EM, Tomassini JE;
 PI WPI; 1985-100818/17.
 DR N-PSDB; AAN50274.
 DR
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus.
 XX
 PS Claim 21; Page 46-48; 49pp; English.
 XX
 CC VP1 is isolated by solubilisation of the intact virus in an aq. anionic
 CC surfactant and a reducing agent. The viral proteins are sepd. and the
 CC protein of molecular wt. 33000 daltons is sepd
 XX
 SQ Sequence 366 AA;
 Query Match 100.0%; Score 103; DB 1; Length 366;
 Best Local Similarity 100.0%; Pred. No. 1.9e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 319 QRLKYAQEELSNEVLPPPRK 338

RESULT 6
AAP50287
ID AAP50287 standard; protein; 854 AA.
AC AAP50287;
XX
XX 25-MAR-2003 (revised)
DT 30-NOV-1991 (first entry)
XX
XX Hepatitis A virus (HAV) peptide corresponding to the capsid protein
DE region of poliovirus RNA.
DE
XX
XX Hepatitis A virus assay; antigen; antibody.
XX
XX Hepatitis A virus.
OS
XX
XX WO8501517-A.
PN
XX
XX 11-APR-1985.
PD
XX
XX 27-SEP-1984; 84WO-US001552.
PF
XX
XX 30-SEP-1983; 83US-00537911.
PR
XX
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
PA
XX
XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
PI Racaniello VR;
FI
XX
XX WPI; 1985-098846/16.
DR
XX
XX N-PSDB; AAN50330.
DR
XX
XX New hepatitis A virus CDNA - useful in assays for the virus and for
PT prodn. of the viral antigen and antibodies to it.
PT
XX
XX Example; Fig 7; 60pp; English.
FS
XX
XX The inventors claim HAV cDNA and a method for producing it, whereby large
CC amts. can be obtd. economically. The cDNA is useful in the assay for
CC detection of HAV quickly and easily and with high sensitivity and
CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
CC 2003 to correct PA field.)
CC
SQ Sequence 854 AA;

Query Match 100.0%; Score 103; DB 1; Length 854;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 7
AAP50116
ID AAP50116 standard; protein; 993 AA.
XX
XX AAP50116;
AC
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-SEP-1991 (first entry)
XX
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
DE and VP-4.

Query Match 100.0%; Score 103; DB 1; Length 854;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 7
AAP50116
ID AAP50116 standard; protein; 993 AA.
XX
XX AAP50116;
AC
XX
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 30-SEP-1991 (first entry)
XX
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
DE and VP-4.

XX Antigenic protein; immunogen; vaccine.
KW
XX
OS Hepatitis A virus; (strain CR326).
XX
XX EP154587-A.
PN
XX
XX 11-SEP-1985.
PD
XX
XX 27-FEB-1985; 85EP-00400369.
PF
XX
XX 02-MAR-1984; 84US-00585818.
FR
XX
XX (MERI) MERCK & CO INC.
PA
XX
XX Linemeyer DL, Menke JG, Rueben RG, Mitra SW;
PI
XX
XX WPI; 1985-224964/37.
DR
XX
XX N-PSDB; AAN50139.
DR
XX
XX New nucleotide sequences coding for hepatitis A virus antigens - useful
PT for eliciting normal immune response and in vaccines for protecting
PT against the virus.
PT
XX
XX Example; Page 11-17; 32pp; English.
PS
XX
XX Within the sequence in AAN50139 is encoded the information necessary to
CC make the antigenic proteins of HAV. The sequences encoding for the
CC structural proteins begin at base 403. The key sub-unit sequences within
CC VP-1, designated Sequences I,II,III,IV, and V, start, respectively at
CC 1882, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
CC valuable as encoding antigenic proteins are the sequences from base 1749
CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
CC 2722. the sequence from base 1749 to base 2722 is esp. valuable as a
CC vector for producing antigen protein. Sequences II-V are claimed. X in
CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
CC
XX
SQ Sequence 993 AA;

Query Match 100.0%; Score 103; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 946 QRLKYAQEELSNEVLPPPRK 965

RESULT 8
AAP50231
ID AAP50231 standard; protein; 993 AA.
XX
XX AAP50231;
AC
XX
XX 28-NOV-1991 (first entry)
DT
XX
XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
DE
XX
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
XX Hepatitis A virus.
OS
XX
XX Key Location/Qualifiers
FH 628..993
FT /note="claimed; X denotes translated stop codons and
FT unspecified triplets"
XX
XX EP138704-A.
PN
XX
XX 24-APR-1985.
PD

XX -type HAV genome in which the P2 region is replaced by the P2 region from
 CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
 CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
 CC sufficient for establishment of infection and accelerated growth in cell
 CC culture
 XX
 SQ Sequence 1077 AA;
 Query Match 100.0%; Score 103; DB 2; Length 1077;
 Best Local Similarity 100.0%; Pred. No. 5.8e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 DB 1033 QRLKYAQEELSNEVLPPPRK 1052
 RESULT 10
 AAR32426
 ID AAR32426 standard; protein; 1091 AA.
 AC AAR32426;
 XX 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-DEC-2001 (revised)
 DT 10-JUN-1993 (first entry)
 XX
 DE Translated from 5' region of Hepatitis A Virus genomic clone.
 XX HAV HM-175; chronic liver disease; picornavirus.
 XX Hepatitis A virus.
 XX
 FH Key Location/Qualifiers
 FT Region 1..711
 FT /note= "X's correspond to nonsense codons, i.e. this
 FT region is not an ORF"
 FT 238..1091
 FT /label= ORF
 FT /note= "second putative initiation codon at position 240"
 XX
 PN USN7788262-N.
 XX
 PD 15-DEC-1992.
 XX
 PF 06-NOV-1991; 91US-00788262.
 XX
 PR 30-SEP-1983; 88US-00536911.
 PR 27-SEP-1984; 84US-00654942.
 PR 06-OCT-1988; 88US-00256135.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Ticehurst JR, Baltimore D, Feinstone SM, Purcell RH;
 PI Racaniello VR, Baroudy BM, Emerson SU;
 XX
 DR WPI; 1993-067429/08.
 DR N-PSDB; AAQ36934.
 XX
 XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
 PT antigen and antibodies.
 XX
 PS Disclosure; Fig 7; 65pp; English.
 XX
 CC HAV virion RNA was extracted from the livers of marmosets which had been
 CC inoculated with HAV (the HAV had previously been passaged twice in
 CC marmosets). The RNA was used to prepare ds cDNA clones by standard
 CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected
 CC African Green Monkey kidney cells were selected for further analysis. A
 CC 7.4kb restriction map (about 99% of the HAV genome) was constructed from
 CC 5 overlapping inserts. The sequence of the first 3.3kb (approx.) from the
 CC 5'-terminus was determined. An amino acid sequence was decoded from the

XX 09-OCT-1984; 84EP-00402025.
 XX
 PR 14-OCT-1983; 83US-00541836.
 PR 02-MAR-1984; 84US-00585942.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Hughes JV, Scolnick EM, Tomassini JE;
 XX
 DR WPI; 1985-100818/17.
 DR N-PSDB; AAN50274.
 XX
 XX New hepatitis A virus surface protein - useful for binding to
 PT neutralising antibodies to the virus.
 XX
 PS Disclosure; Page 17-23; 49pp; English.
 XX
 CC VP1 is isolated by solubilisation of the intact virus in an aq. anionic
 CC surfactant and a reducing agent. The viral proteins are sepd. and the
 CC protein of molecular wt. 33000 daltons is sepd
 XX
 SQ Sequence 993 AA;
 Query Match 100.0%; Score 103; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 5.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 DB 946 QRLKYAQEELSNEVLPPPRK 965
 RESULT 9
 AAW95559
 ID AAW95559 standard; protein; 1077 AA.
 AC AAW95559;
 DT 28-APR-1999 (first entry)
 XX
 DE A partial hepatitis A virus (HAV) protein.
 XX
 KW Hepatitis A virus protein; HAV; P2 region;
 KW cell-culture-adapted HAV strain; infection; accelerated growth.
 XX
 OS Hepatitis A virus.
 XX
 PN US5849562-A.
 XX
 PD 15-DEC-1998.
 XX
 PF 06-JUN-1995; 95US-00468926.
 XX
 PR 30-SEP-1983; 83US-00537911.
 PR 27-SEP-1984; 84US-00654942.
 PR 06-OCT-1988; 88US-00256135.
 PR 06-NOV-1991; 91US-00788262.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Emerson SU, Purcell RH;
 XX
 DR WPI; 1999-094412/08.
 DR N-PSDB; AAX01006.
 XX
 XX Chimeric hepatitis A virus strains - with P2 region from cell-culture-
 PT adapted strain in wild-type genome.
 XX
 PS Disclosure; Fig 7A-L; 36pp; English.
 XX
 CC The present sequence represents a partial hepatitis A virus (HAV)
 CC protein. The specification describes a DNA construct consisting of a wild

PR 19-SEP-1984; 84US-00652067.
 PR 09-SEP-1986; 86US-00905146.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
 PI Daemer RJ, Gust ID;
 XX WPI; 1990-075557/10.
 DR N-PSDB; RAQ03512.
 XX Vaccine against hepatitis A virus infection - comprises novel attenuated
 FT hepatitis A virus strain.
 PT Claim 1; Fig 1; 18pp; English.
 XX The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
 CC standardise OS field)
 XX Sequence 2227 AA;
 SQ
 Query Match 100.0%; Score 103; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q/ 1 QRLKYAQEELSNEVLPPPRK 20
 Db 810 QRLKYAQEELSNEVLPPPRK 829
 RESULT 13
 AAW34074
 ID AAW34074 standard; protein; 2227 AA.
 XX AC AAW34074;
 XX 17-OCT-2003 (revised)
 DT 27-APR-1998 (first entry)
 XX Hepatitis A virus HM-175 protein sequence.
 XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
 KW vaccine.
 XX Hepatitis A virus; HM-175.
 OS
 XX Key Location/Qualifiers
 FT Protein 1..23
 FT /label= VP4
 FT Protein 24..245
 FT /label= VP2
 FT Protein 246..491
 FT /label= VP3
 FT Protein 492..791
 FT /label= VP1
 FT Protein 792..980
 FT /label= 2A
 FT Protein 981..1087
 FT /label= 2B
 FT Protein 1088..1422
 FT /label= 2C
 FT Protein 1423..1496
 FT /label= 3A
 FT Protein 1497..1519
 FT /label= 3B
 FT Protein 1520..1738

FT Protein /label= 3C
 FT 1739..2227
 FT /label= 3D
 PN WO9740166-A2.
 XX 30-OCT-1997.
 XX 18-APR-1997; 97WO-US006506.
 PF 19-APR-1996; 96US-0015642P.
 PR (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Raychaudhuri G, Emerson SU, Purcell RH;
 PI WPI; 1997-535850/49.
 XX N-PSDB; AAT93023.
 DR Human attenuated HAV genome containing simian HAV 2C gene - useful as
 XX vaccines against HAV infection.
 PT Disclosure; Fig 13A-D; 66pp; English.
 PS This protein sequence is encoded by the human hepatitis A virus (HAV) HM-
 CC 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained
 CC by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA
 CC construct (1) comprises a genome of HAV, where the genome is a human
 CC attenuated HAV genome in which a region of the 2C gene has been replaced
 CC by a corresponding region from a 2C gene of a simian AGM-27 HAV genome
 CC (see AAT93024). The region of the 2C gene from AGM-27 contained in the
 CC construct preferably encodes amino acids 120-328 of the 2C protein, amino
 CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript
 CC of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3)
 CC a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host
 CC cell containing the HAV of (3). (1) or its RNA transcript, can be used as
 CC a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can
 CC also be used to stimulate the production of protective antibodies in the
 CC mammal. (Updated on 17-OCT-2003 to standardise OS field)
 XX SQ Sequence 2227 AA;
 Query Match 100.0%; Score 103; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Q/ 1 QRLKYAQEELSNEVLPPPRK 20
 Db 810 QRLKYAQEELSNEVLPPPRK 829
 RESULT 14
 AAB18609
 ID AAB18609 standard; protein; 2227 AA.
 XX AC AAB18609;
 XX 15-JAN-2001 (first entry)
 DT Amino acid sequence of live attenuated Hepatitis A virus 4380.
 DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
 XX HAV 4380.
 KW Hepatitis A virus.
 XX US6113912-A.
 PN 05-SEP-2000.
 PD 07-JUN-1995; 95US-00475886.
 XX 18-SEP-1992; 92US-00947338.
 PR

PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75476.
 XX
 XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.
 XX
 XX Disclosure; Col 93-104; 72pp; English.
 XX
 XX The present sequence is derived from a live attenuated hepatitis A virus
 CC (HAV) of the invention, designated HAV 4380. The sequence is produced by
 CC modifying wild type HAV strain HM-174. The HAV of the invention are
 CC adapted to growth in the human fibroblast-like cell line MRC-5. The HAV
 CC is able to propagate in MRC-5 cells and retain appropriate attenuation.
 CC It is useful as a live vaccine for prophylaxis of hepatitis A in humans
 CC and other primates
 XX
 XX SQ Sequence 2227 AA;
 XX
 XX Query Match 100.0%; Score 103; DB 3; Length 2227;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 |||||
 DB 810 QRLKYAQEELSNEVLPPPRK 829
 |||||
 RESULT 15
 AAB18607
 ID AAB18607 standard; protein; 2227 AA.
 AC
 AC AAB18607;
 XX
 XX 15-JAN-2001 (first entry)
 DT
 XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.
 DE
 XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.
 KW
 XX Hepatitis A virus.
 OS
 XX US6113912-A.
 PN
 XX 05-SEP-2000.
 PD
 XX 07-JUN-1995; 95US-00475886.
 PF
 XX 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 XX
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
 XX
 XX WPI; 2000-586464/55.
 DR N-PSDB; AAA75476.
 XX
 XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
 PT line useful as vaccine for protecting humans against hepatitis A virus
 PT infection, has modified genome compared to wild type.
 XX
 XX Disclosure; Fig 6A-K; 72pp; English.
 PS
 XX The present sequence is derived from a wild type hepatitis A virus (HAV)

CC strain HM-174. The sequence is modified to produce HAV which are adapted
 CC to growth in the human fibroblast-like cell line MRC-5. The HAV is able
 CC to propagate in MRC-5 cells and retain appropriate attenuation. It is
 CC useful as a live vaccine for prophylaxis of hepatitis A in humans and
 CC other primates
 XX
 XX SQ Sequence 2227 AA;
 XX
 XX Query Match 100.0%; Score 103; DB 3; Length 2227;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 |||||
 DB 810 QRLKYAQEELSNEVLPPPRK 829
 |||||
 RESULT 16
 AAB19899
 ID AAB19899 standard; protein; 2227 AA.
 XX
 XX AAE19899;
 AC
 XX 18-JUN-2002 (first entry)
 DT
 XX Hepatitis A virus (HAV) protein.
 DE
 XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
 XX
 XX Hepatitis A virus.
 OS
 XX WO200213855-A2.
 PN
 XX 21-FEB-2002.
 PD
 XX 15-AUG-2001; 2001WO-IB001808.
 PF
 XX 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 PR 03-NOV-2000; 2000US-00705547.
 XX
 XX (TRIP-) TRIPEP AB.
 PA
 XX Sallberg M, Hultgren C;
 PI
 XX WPI; 2002-241837/29.
 DR N-PSDB; AAD31766.
 DR
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 PT
 XX Claim 11; Page 82-87; 120pp; English.
 PS
 XX The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as
 CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is hepatitis A virus (HAV) protein
 XX
 XX SQ Sequence 2227 AA;
 XX
 XX Query Match 100.0%; Score 103; DB 5; Length 2227;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e-06;
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 |||||
 DB 810 QRLKYAQEELSNEVLPPPRK 829
 |||||

```

RESULT 17
ABG31729
ID  ABG31729 standard; protein; 2227 AA.
XX  AC
XX  ABG31729;
XX  AC
XX  DT 29-AUG-2003 (revised)
XX  DT 29-NOV-2002 (first entry)
XX  DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX  DE Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX  KW HAV 4380.
XX  OS
XX  OS Hepatitis A virus; strain HM-175.
XX  PN
XX  PN US6423318-B1.
XX  PD
XX  PD 23-JUL-2002.
XX  PF
XX  PF 31-AUG-2000; 2000US-00653499.
XX  PR
XX  PR 17-SEP-1993; 93WO-US008610.
XX  PR 17-APR-1995; 95US-00397232.
XX  PR 07-JUN-1995; 95US-00475886.
XX  PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX  PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX  PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX  WPI; 2002-680946/73.
XX  DR N-PSDB; ABS52789.
XX  DR
XX  PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX  PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX  PS Disclosure; Col 93-104; 71pp; English.
XX  CC The invention relates to a polynucleotide which encodes a hepatitis A
XX  CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX  CC line). The polynucleotide is useful for preparing a vaccine against
XX  CC hepatitis A virus infection. This sequence represents an attenuated
XX  CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
XX  CC standardise OS field)
XX  SQ Sequence 2227 AA;
XX  Query Match 100.0%; Score 103; DB 5; Length 2227;
XX  Best Local Similarity 100.0%; Pred. No. 1.2e-06;
XX  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX  QY 1 QRLKYAQEELSNEVLPPPRK 20
XX  DB 810 QRLKYAQEELSNEVLPPPRK 829
XX  RESULT 18
ABG31727
ID  ABG31727 standard; protein; 2227 AA.
XX  AC
XX  AC ABG31727;
XX  DT 29-AUG-2003 (revised)
XX  DT 29-NOV-2002 (first entry)
XX  DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX  DE Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX  KW Hepatitis A virus; strain HM-175.
XX  OS Hepatitis A virus; strain HM-175.
```

```

XX  US6423318-B1.
XX  23-JUL-2002.
XX  31-AUG-2000; 2000US-00653499.
XX  17-SEP-1993; 93WO-US008610.
XX  17-APR-1995; 95US-00397232.
XX  07-JUN-1995; 95US-00475886.
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX  (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX  Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX  WPI; 2002-680946/73.
XX  DR N-PSDB; ABS52787.
XX  DR
XX  PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX  PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
XX  PS Disclosure; Fig 6; 71pp; English.
XX  CC The invention relates to a polynucleotide which encodes a hepatitis A
XX  CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX  CC line). The polynucleotide is useful for preparing a vaccine against
XX  CC hepatitis A virus infection. This sequence represents a hepatitis A virus
XX  CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
XX  CC field)
XX  SQ Sequence 2227 AA;
XX  Query Match 100.0%; Score 103; DB 5; Length 2227;
XX  Best Local Similarity 100.0%; Pred. No. 1.2e-06;
XX  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX  QY 1 QRLKYAQEELSNEVLPPPRK 20
XX  DB 810 QRLKYAQEELSNEVLPPPRK 829
XX  RESULT 19
ABU08641
ID  ABU08641 standard; protein; 2227 AA.
XX  AC
XX  AC ABU08641;
XX  DT 23-OCT-2003 (revised)
XX  DT 03-JUN-2003 (first entry)
XX  DE Attenuated hepatitis A virus (4380) strain HM-175.
XX  KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX  KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX  OS
XX  OS Hepatitis A virus; strain HM-175.
XX  PN
XX  PN US2002176869-A1.
XX  PD
XX  PD 28-NOV-2002.
XX  PF
XX  PF 29-APR-2002; 2002US-00135988.
XX  PR
XX  PR 18-SEP-1992; 92US-00947338.
XX  PR 17-SEP-1993; 93WO-US008610.
XX  PR 17-APR-1995; 95US-00397232.
XX  PR 07-JUN-1995; 95US-00475886.
XX  PR 31-AUG-2000; 2000US-00653499.
XX  PA (FUNK/) FUNKHOUSER A W.
XX  PA (EMER/) EMERSON S U.
XX  PA (PURC/) PURCELL R H.
```

PA (DHON/) D'HONDT E.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93475.
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX Disclosure; Page 45-51; 70pp; English.
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated human
CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX Sequence 2227 AA;
SQ Query Match 100.0%; Score 103; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829
RESULT 20
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX AC ABU08639;
XX 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX Wild type human hepatitis A virus strain HM-175.
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX vaccine; MRC-5 cell; hepatitis infection.
XX Hepatitis A virus; strain HM-175.
OS US2002176869-A1.
XX FN 28-NOV-2002.
XX 29-APR-2002; 2002US-00135988.
XX 18-SEP-1992; 92US-00947338.
PR 17-SEP-1993; 93WO-US008610.
PR 17-APR-1995; 95US-00397232.
PR 07-JUN-1995; 95US-00475886.
PR 31-AUG-2000; 2000US-00653499.
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93473.
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX

PS Disclosure; Fig 6; 70pp; English.
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)
XX Sequence 2227 AA;
SQ Query Match 100.0%; Score 103; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829
RESULT 21
ABW00350
ID ABW00350 standard; protein; 2227 AA.
XX AC ABW00350;
XX 15-JAN-2004 (first entry)
DT Hepatitis A virus protein.
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide.
XX Hepatitis A virus.
XX US2002136740-A1.
XX 26-SEP-2002.
XX 15-AUG-2001; 2001US-00929955.
XX 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
XX (SALL/) SALLBERG M.
PA (HULT/) HULTGREN C.
XX Sallberg M, Hultgren C;
XX WPI; 2003-764978/72.
DR N-PSDB; AAD60867.
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX Claim 11; Page 45-51; Opp; English.
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis A
CC virus protein
XX Sequence 2227 AA;
SQ Query Match 100.0%; Score 103; DB 7; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 20

Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 22

AA18608

ID AAB18608 standard; protein; 2227 AA.

XX AC AAB18608;

XX DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX KW P-35 virus.

XX OS Hepatitis A virus.

XX PN US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-00475886.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US008610.

XX PR 17-APR-1995; 95US-00397232.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

DR N-PSDB; AAA75477.

XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

PT infection, has modified genome compared to wild type.

XX Disclosure; Col 67-78; 72pp; English.

XX CC The present sequence is derived from passage 35 of a wild type hepatitis

CC A virus (HAV) strain HM-174. The resulting virus is designated P-35

CC virus. The sequence is modified to produce HAV which are adapted to

CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to

CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful

CC as a live vaccine for prophylaxis of hepatitis A in humans and other

CC primates

XX SQ Sequence 2227 AA;

Query Match 95.1%; Score 98; DB 3; Length 2227;

Best Local Similarity 95.0%; Pred. No. 7.2e-06;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

|||||:|||||

Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 23

ABG31728

ID ABG31728 standard; protein; 2227 AA.

XX AC ABG31728;

XX DT 29-NOV-2002 (first entry)

XX DE Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.

XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast;

XX KW virucide; mutant; pHAV/7; mutain.

OS Hepatitis A virus; strain HM-175.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 764 /note= "Wild-type Glu substituted by Val"

FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"

FT Misc-difference 963 /label= Wild-type Lys substituted by Arg

FT Misc-difference 1052 /note= "Wild-type Ala substituted by Val"

FT Misc-difference 1062 /note= "Wild-type Gly substituted by Ala"

FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"

FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"

FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"

FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"

FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"

FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"

FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"

XX US64233318-B1.

XX 23-JUL-2002.

XX 31-AUG-2000; 2000US-00653499.

XX 17-SEP-1993; 93WO-US008610.

XX 17-APR-1995; 95US-00397232.

XX 07-JUN-1995; 95US-00475886.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.

XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;

XX WPI; 2002-680946/73.

XX N-PSDB; ABS52788.

XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

XX in MRC-5 cells, useful for preparing a vaccine against HAV infection.

XX Example 3; Col 67-78; 71pp; English.

XX The invention relates to a polynucleotide which encodes a hepatitis A

XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell

XX line). The polynucleotide is useful for preparing a vaccine against

XX hepatitis A virus infection. This sequence represents a hepatitis A virus

XX mutant strain HM-175/7 (pHAV/7) polypeptide

XX Sequence 2227 AA;

Query Match 95.1%; Score 98; DB 5; Length 2227;

Best Local Similarity 95.0%; Pred. No. 7.2e-06;

Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20

|||||:|||||

Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 24

ABU08640

ID ABU08640 standard; protein; 2227 AA.

XX

AC ABU08640;
XX
XX 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
XX Attenuated (pass35) hepatitis A virus strain HM-175.
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
XX
XX Hepatitis A virus; strain HM-175.
XX
XX US2002176869-A1.
XX
XX 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX 07-JUN-1995; 95US-00475886.
XX 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
PI
XX WPI; 2003-352605/02.
XX N-PSDB; ABX93474.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
FT and disease.
XX
XX Example 3; Fig 6; 70pp; English.
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated (pass 35)
CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 2227 AA;
SQ
Query Match 95.1%; Score 98; DB 6; Length 2227;
Best Local Similarity 95.0%; Pred. No. 7.2e-06;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829
RESULT 25
AAR15629
ID AAR15629 standard; protein; 839 AA.
XX
XX AAR15629;
XX
XX 24-OCT-2003 (revised)
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 17-MAR-1992 (first entry)
XX
XX Capsid region of cyno-HAV isolate CY-145.
DE
XX Hepatitis A virus; cynomolgus; HAV; monkey; vaccine; macaque.
KW

XX Viruses.
XX
XX Key Location/Qualifiers
FT Protein 1..245
FT Cleavage-site /label= VP0
FT Protein 245..246
FT Cleavage-site 246..491
FT Active-site /label= VP3
FT Cleavage-site 315
FT Cleavage-site 491..492
FT Protein 492..791
FT Active-site /label= VP1
FT Cleavage-site 593
FT Cleavage-site 791..792
FT Protein 792
FT /label= P2
FT /note= "incomplete"
XX
XX USN7678828-N.
PN
XX 12-NOV-1991.
PD
XX 03-APR-1991; 91US-00678828.
PF
XX 03-APR-1991; 91US-00678828.
PR
XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
PA
XX Hainan OV, Margolis HS, Robertson BH, Brinton MH, Ebert JW;
PI
XX WPI; 1991-376737/51.
XX N-PSDB; AAQ15180.
DR
XX
XX Hepatitis A virus isolates and DNA - used to prepare vaccines for
PT preventing hepatitis A virus infection.
XX
XX Disclosure; Fig 3; 23pp; English.
XX
XX The sequence was deduced from the nucleotide sequence obtd. by PCR
CC amplification of cyno-HAV viral RNA obtd. from the stool of a cynomolgus
CC monkey with serologically and histologically confirmed spontaneous
CC hepatitis A. The sequence differs from the human HAV isolate HMI75
CC (Cohen, J.I., et al. (1987) Proc. Natl. Acad. Sci. USA 84, 2497-2501),
CC mainly in the VP3 and VP1 proteins. The Gln-Val pair at the VP3-VP1
CC cleavage site in the human isolate is replaced by a Gln-Thr pair in the
CC cyno-HAV. The other two cleavage sites are the same. Two residues have
CC been identified as part of the immuno- dominant region (see feature
CC table) and are different to those in the same position in human HAV. The
CC protein and peptides derived from it can be used in the prepn. of
CC vaccines for the prevention of HAV infection. See also AAR15056. (Note:
CC Revised entry submitted to correct the patent number format of US
CC Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis.us.html.) (Updated on 25-
CC MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 839 AA;
SQ
Query Match 91.3%; Score 94; DB 2; Length 839;
Best Local Similarity 85.0%; Pred. No. 1.1e-05;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 809 QRLKYAQEELSNEVLPPPRK 828
RESULT 26
ABB68507
ID ABB68507 standard; protein; 150 AA.
XX


```

AB 68507;
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32313.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX
XX WPI; 2001-556860/75.
DR
XX N-PSDB; ABL12610.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Disclosure; SEQ ID NO 32313; 21pp + Sequence Listing; English.
PS
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
XX ABB72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 150 AA;
SQ
Query Match 48.5%; Score 50; DB 4; Length 150;
Best Local Similarity 52.4%; Pred. NO. 10;
Matches 11; Conservative 3; Mismatches 3; Indels 4; Gaps 1;
QY 1 ORLKVAQ-----BELSNEVLPP 17
DB 32 RELKYRQRDRVSEIANEYLPP 52
RESULT 27
AAW20624
ID AAW20624 standard; protein; 444 AA.
XX
XX AAW20624;
AC
XX
XX 14-JUL-1997 (first entry)
DT
XX
XX H. pylori cytoplasmic protein, 02aell612orf25.
DE
XX
XX Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacterium; life cycle; activator;
KW bacteria; inhibitor; duodenal ulcer disease; chronic gastritis;
KW diagnosis.
XX
XX Helicobacter pylori.
OS
XX
XX WO9640893-A1.
PN

```

```

XX
PD 19-DEC-1996.
XX
XX 06-JUN-1996; 96WO-US009122.
XX
XX 07-JUN-1995; 95US-00487032.
PR
XX 01-APR-1996; 96US-00630405.
XX
XX (ASTR ) ASTRA AB.
PA
XX
XX Smith D, Berglindh OT, Mellgaard BL;
PI
XX
XX WPI; 1997-052306/05.
DR
XX N-PSDB; AAT67877.
XX
XX Helicobacter pylori nucleic acid sequences and related polypeptide(s) -
PT useful for vaccines to treat or prevent H. pylori infection, and to
PT detect Helicobacter.
XX
XX Claim 61; Page 1050-51; 1481pp; English.
PS
XX
XX This sequence represents a H. pylori cytoplasmic protein. The protein may
XX be used in a vaccine to prevent or treat H. pylori infection or to
XX identify H. pylori polypeptide binding compounds, useful as potential H.
XX pylori life cycle activators or inhibitors. The genomic sequence of H.
XX pylori (ATCC 55679) was determined from overlapping contigs generated by
XX mechanically shearing the bacterial DNA. The sequences were analysed for
XX ORF of at least 180 nucleotides, and the predicted coding regions defined
XX by computer evaluation. To identify likely H. pylori antigens for vaccine
XX development, the amino acid sequences predicted from various ORF were
XX analysed for significant homology to other known or exported membrane
XX proteins. Having identified and determined the sequences of interest,
XX particular regions can be isolated from H. pylori by PCR amplification
XX for recombinant polypeptide production, e.g. in E. coli hosts
XX
XX Sequence 444 AA;
SQ
Query Match 46.6%; Score 48; DB 2; Length 444;
Best Local Similarity 75.0%; Pred. NO. 64;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 ELSNEVLPPPRK 20
DB 416 EADNEELPPPRK 427
RESULT 28
ABB59692
ID ABB59692 standard; protein; 1176 AA.
XX
XX ABB59692;
AC
XX
XX 26-MAR-2002 (first entry)
DT
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 5868.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
XX Drosophila melanogaster.
OS
XX
XX WO200171042-A2.
PN
XX
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI

```

XX WPI; 2001-656860/75.
DR N-PSDB; ABL03795.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 5868; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1176 AA;
Query Match 46.6%; Score 48; DB 4; Length 1176;
Best Local Similarity 56.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPP 16
:|||||:|:|:
DB 335 KRKYAKELQKEMLP 350
RESULT 29
ABB68504
ID ABB68504 standard; protein; 221 AA.
XX
AC ABB68504;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 32304.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR N-PSDB; ABL12607.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 32304; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 221 AA;
Query Match 45.6%; Score 47; DB 4; Length 221;
Best Local Similarity 47.6%; Pred. No. 44;
Matches 10; Conservative 4; Mismatches 3; Indels 4; Gaps 1;
QY 1 QRLKYAQEELSNEVLPP 17
:|||||:|:|:
DB 128 RLKRYQRQDVSDIANEVLPP 148
RESULT 30
AAW42923
ID AAW42923 standard; peptide; 20 AA.
XX
AC AAW42923;
XX
DT 28-APR-1998 (first entry)
XX
DE Immunogenic Hepatitis A virus peptide YK-1316.
XX
KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
KW antibody.
XX
OS Synthetic.
OS Hepatitis A virus.
XX
PN WO9740147-A1.
XX
PD 30-OCT-1997.
XX
PF 18-APR-1997; 97WO-US006891.
XX
PR 19-APR-1996; 96US-0015644P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Fields HA, Khudyakov YE;
XX
DR WPI; 1997-535831/49.
XX
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.
XX
PS Claim 18; Page 112; 140pp; English.
XX
CC Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. The present peptide
CC is derived from amino acids 799-818, and has a reactivity of 41.7% with
CC acute sera. Compositions containing the peptides can be used to induce an
CC immune response to HAV in a mammal. The peptides can also be used to
CC detect the presence of antibodies against HAV in mammalian serum. The
CC peptides can also be used to make an antibody against HAV by
CC administering the peptide to a mammal
XX
SQ Sequence 20 AA;
Query Match 43.7%; Score 45; DB 2;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEE 9
:|||||:

Db 12 QRLKYAEE 20

Search completed: May 11, 2004, 13:37:28
Job time : 50 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:30:01 ; Search time 10.5 Seconds

(without alignments)

183.222 Million cell updates/sec

Title: US-09-171-432A-41
Perfect score: 103
Sequence: 1 ORLYAQEELSNEVLPPPK 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR 78:.*
1: Pir1:.*
2: Pir2:.*
3: Pir3:.*
4: Pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	103	100.0	341	2	S04137 genome polypeptide
2	103	100.0	852	1	GNNYHA genome polypeptide
3	103	100.0	1358	2	A03905 genome polypeptide
4	103	100.0	2227	1	GNNYHM genome polypeptide
5	103	100.0	2227	1	GNNYHB genome polypeptide
6	103	100.0	2227	1	GNNYMK genome polypeptide
7	98	95.1	2227	1	GNNYSA genome polypeptide
8	97	94.2	2230	1	GNNYS2 genome polypeptide
9	94	91.3	839	1	GNNYS2 regulatory protein
10	56	54.4	346	2	S74448 DNA-directed RNA p
11	52	50.5	859	2	T43701 protein C26E6.4 [i
12	52	50.5	1193	2	E88445 related to cytochrome
13	51	49.5	1119	2	T50995 proteinase HslVU
14	49.5	48.1	443	2	E82046 genome polypeptide
15	49	47.6	55	2	PQ0433 genome polypeptide
16	49	47.6	56	2	PQ0434 genome polypeptide
17	49	47.6	56	2	PQ0428 genome polypeptide
18	49	47.6	56	2	PQ0427 genome polypeptide
19	49	47.6	56	2	PQ0429 genome polypeptide
20	49	47.6	56	2	PQ0432 genome polypeptide
21	49	47.6	56	2	PQ0430 genome polypeptide
22	49	47.6	1174	2	T28976 DNA-directed RNA p
23	48	46.6	442	2	A71969 probable histidine
24	48	46.6	1176	2	A27826 DNA-directed RNA p
25	48	46.6	1191	2	S65068 DNA-directed RNA p
26	47	45.6	592	1	LLBY actin-binding prot
27	47	45.6	6642	2	T29757 protein UNC-89 - C
28	46	44.7	338	2	I56893 transcription fact
29	46	44.7	480	1	FWPU1B 11S globulin beta

30	46	44.7	483	2	T01053 hypothetical prote
31	45	43.7	756	2	JC5590 cycloartenol synth
32	45	43.7	1188	2	T05846 DNA-directed RNA p
33	45	43.7	2420	2	A84652 hypothetical prote
34	44.5	43.2	1609	2	T01797 hypothetical prote
35	44	42.7	56	2	PQ0436 genome polypeptide
36	44	42.7	223	2	PQ0514 hemagglutinin HA2
37	44	42.7	303	2	S61723 phosphoglycerate m
38	44	42.7	443	2	AH0013 Arp-binding heat s
39	44	42.7	510	2	S42705 probable phosphogl
40	44	42.7	569	2	T43531 probable potassium
41	44	42.7	694	2	T10565 hypothetical prote
42	44	42.7	914	2	S48333 ORC1 protein - yea
43	43.5	42.2	361	2	G82530 hypothetical prote
44	43.5	42.2	3678	2	S28916 dystrophin - mouse
45	43.5	42.2	3685	1	A27605 dystrophin, muscle
46	43	41.7	36	2	S78721 protein YGL006w-a
47	43	41.7	56	2	PQ0437 genome polypeptide
48	43	41.7	120	2	D69679 alkaline-shock pro
49	43	41.7	163	2	PC4186 hemagglutinin 2 ch
50	43	41.7	268	2	S48942 hypothetical prote
51	43	41.7	277	2	T11680 hypothetical prote
52	43	41.7	496	2	G64527 hypothetical prote
53	43	41.7	574	1	HMIVEM hemagglutinin prec
54	43	41.7	576	2	S03300 hemagglutinin prec
55	43	41.7	577	2	A97195 uncharacterized AB
56	43	41.7	578	2	S03299 hemagglutinin prec
57	43	41.7	578	2	S03301 hemagglutinin prec
58	43	41.7	583	1	HMIVBS hemagglutinin prec
59	43	41.7	583	1	HMIVHO hemagglutinin prec
60	43	41.7	585	1	HMIVBJ hemagglutinin prec
61	43	41.7	727	2	AB2353 hypothetical prote
62	43	41.7	763	2	B86250 hypothetical prote
63	43	41.7	869	2	H89864 hypothetical prote
64	43	41.7	870	2	E72541 probable aconitate
65	43	41.7	1126	2	A96032 probable two-compo
66	43	41.7	1144	2	A75132 hypothetical prote
67	43	41.7	1146	2	S64402 probable transcrip
68	43	41.7	1210	2	S35548 DNA-directed RNA p
69	43	41.7	1345	2	B71508 DNA-directed RNA p
70	43	41.7	2588	2	T14342 MSD1 protein - mou
71	42.5	41.3	204	2	C97389 hypothetical prote
72	42.5	41.3	204	2	AD2607 conserved hypothetical
73	42.5	41.3	465	2	H96751 probable casein ki
74	42.5	41.3	2278	1	S56274 FABI protein - yea
75	42	40.8	201	2	C96813 hypothetical prote
76	42	40.8	214	2	C49941 devs protein - Myx
77	42	40.8	265	2	B85715 hypothetical prote
78	42	40.8	276	2	S56550 hypothetical 30.6X
79	42	40.8	319	2	G85818 hypothetical prote
80	42	40.8	351	2	T29369 hypothetical prote
81	42	40.8	368	2	H90998 probable proteinase
82	42	40.8	441	2	F84560 purple acid phosph
83	42	40.8	445	2	I38027 carl protein - h
84	42	40.8	554	2	A56730 carl protein - pod
85	42	40.8	657	2	H71422 hypothetical prote
86	42	40.8	937	1	DUVZ2W DNA-directed DNA p
87	42	40.8	937	2	A24878 DNA-directed DNA p
88	42	40.8	938	2	T46924 probable translati
89	42	40.8	1005	2	C36842 E9L protein - vari
90	42	40.8	1005	2	H72156 C9L protein - vari
91	42	40.8	1005	2	T28488 probable DNA-direc
92	42	40.8	1006	1	DJVZ41 DNA-directed DNA p
93	42	40.8	1006	2	T30802 DNA-directed DNA p
94	42	40.8	1066	2	T30903 arachidonate 8-lip
95	42	40.8	1209	1	E70627 hypothetical prote
96	42	40.8	1279	2	A47363 RNA helicase A - h
97	42	40.8	1336	2	T18288 ABC transport prot
98	42	40.8	1536	2	S59841 4-alpha-glucanotra
99	41.5	40.3	213	2	T01715 hypothetical prote
100	41.5	40.3	285	2	G85016 probable myb-relat

ALIGNMENTS

RESULT 1
S04137
genome polyprotein - human hepatitis A virus (strain LCDG-1) (fragment)
C:Species: human hepatitis A virus
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C:Accession: S04137
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
Nucleic Acids Res. 17, 3594, 1989
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
A:Reference number: S04137; MUID:89263805; PMID:2542903
A:Accession: S04137
A:Molecule type: mRNA
A:Residues: 1-341 <AND>
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576
C:Genetics:
A:Gene: VP1
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; polyprotein
F:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>

Query Match 100.0%; Score 103; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 314 QRLKYAQEELSNEVLPPPRK 333

RESULT 2
GNVYHA
genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03904
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.K.
J. Virol. 54, 247-255, 1985
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:Reference number: A03904; MUID:85185648; PMID:2985793
A:Accession: A03904
A:Molecule type: genomic RNA
A:Residues: 1-852 <LIN>
A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 103; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 3
A03905
genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein
C:Species: human hepatitis A virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C:Accession: A03905
R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstone
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:Reference number: A03905; MUID:85166289; PMID:2984684
A:Accession: A03905
A:Molecule type: genomic RNA
A:Residues: 1-1358 <BAR>
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein
F:1-245/Product: coat protein 1A #status predicted <C1A>
F:246-491/Product: coat protein 1B #status predicted <C1B>
F:492-836/Product: coat protein 1C #status predicted <C1C>
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 103; DB 2; Length 1358;
Best Local Similarity 100.0%; Pred. No. 8.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 4
GNVYHM
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro
B:RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999
C:Accession: A25981
R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di
A:Reference number: A25981; MUID:87061253; PMID:3023706
A:Accession: A25981
A:Molecule type: genomic RNA
A:Residues: 1-2227 <CH>
A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:1-23/Product: coat protein 1A #status predicted <VP4>
F:24-245/Product: coat protein 1B #status predicted <VP2>
F:246-491/Product: coat protein 1C #status predicted <VP3>
F:492-791/Product: coat protein 1D #status predicted <VP1>
F:792-980/Product: core protein 2A #status predicted <C2A>
F:981-1087/Product: core protein 2B #status predicted <C2B>
F:1088-1422/Product: core protein 2C #status predicted <C2C>
F:1423-1496/Product: protein 3A #status predicted <C3A>
F:1497-1519/Product: protein 3B #status predicted <C3B>
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 103; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-07; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 5
GNVYHR
genome polyprotein - human hepatitis A virus
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA polymerase (EC 2.7.7.48), protein 3D
C:Species: human hepatitis A virus
A:Note: host Homo sapiens (man)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
C:Accession: A03903
R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A;Title: Primary structure and gene organization of human hepatitis A virus.
A;Reference number: A03903; MUID:85190549; PMID:2986127
A;Accession: A03903
A;Molecule type: genomic RNA
A;Residues: 1-2227 <RNA>
A;Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F;1-245/Product: coat protein 1A #status predicted <C1A>
F;245-491/Product: coat protein 1B #status predicted <C1B>
F;492-836/Product: coat protein 1C #status predicted <C1C>
F;837-980/Product: coat protein 2A #status predicted <C2A>
F;981-1076/Product: core protein 2B #status predicted <C2B>
F;1077-1422/Product: core protein 2C #status predicted <C2C>
F;1423-1484/Product: protein 3A #status predicted <C3A>
F;1485-1507/Product: protein 3B #status predicted <C3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 103; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQBELSNEVLPPPRK 20
Db 810 QRLKYAQBELSNEVLPPPRK 829

RESULT 6
GNVYH
Genome polyprotein - human hepatitis A virus (strain MBB)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Accession: J50303
C;Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 10-May-1996
R;Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Kiehn, R.; Wimmer, B.; Deinhardt, Virus Res. 8, 153-171, 1987
A;Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat
A;Reference number: J50303; MUID:88045071; PMID:2823500
A;Accession: J50303
A;Molecule type: genomic RNA
A;Residues: 1-2227 <PAU>
A;Cross-references: EMBL:M20273
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrolase
F;1-23/Product: coat protein 1A #status predicted <VP4>
F;24-246/Product: coat protein 1B #status predicted <VP2>
F;247-491/Product: coat protein 1C #status predicted <VP3>
F;492-836/Product: coat protein 1D #status predicted <VP1>
F;837-980/Product: core protein 2A #status predicted <P2A>
F;981-1108/Product: core protein 2B #status predicted <P2B>
F;1109-1438/Product: core protein 2C #status predicted <P2C>
F;1439-1496/Product: protein 3A #status predicted <P3A>
F;1497-1519/Product: genome-linked protein VPg #status predicted <VPG>
F;1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 103; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQBELSNEVLPPPRK 20
Db 810 QRLKYAQBELSNEVLPPPRK 829

RESULT 7
GNVYH
Genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B; RNA polymerase (EC 2.7.7.48), protein 3D

C;Species: human hepatitis A virus
A;Note: host Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 16-Jul-1999
C;Accession: A94149; A25914; A94508
R;Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R. Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987
A;Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with
A;Reference number: A94149; MUID:87175701; PMID:3031686
A;Accession: A94149
A;Status: nucleic acid sequence not shown
A;Molecule type: genomic RNA
A;Residues: 1-2227 <COH>
A;Cross-references: EMBL:M16632; NID:G329594; PIDN:AAA45471.1; PID:G329595
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase
F;1-245/Product: coat protein 1A #status predicted <P1A>
F;246-491/Product: coat protein 1B #status predicted <P1B>
F;492-836/Product: coat protein 1C #status predicted <P1C>
F;837-980/Product: core protein 2A #status predicted <P2A>
F;981-1076/Product: core protein 2B #status predicted <P2B>
F;1077-1422/Product: core protein 2C #status predicted <P2C>
F;1423-1484/Product: protein 3A #status predicted <P3A>
F;1485-1507/Product: protein 3B #status predicted <P3B>
F;1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>
F;1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 95.1%; Score 98; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 9.2e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQBELSNEVLPPPRK 20
Db 810 QRLKYAQBELSNEVLPPPRK 829

RESULT 8
GNVYH
Genome polyprotein - simian hepatitis A virus (strain AGM-27)
N;Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 2A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C;Species: simian hepatitis A virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jun-2000
C;Accession: A30470; S04885; S03965
R;Tsarev, S.A.
A;Note: submitted to JIPID, April 1991
A;Reference number: A30470
A;Accession: A30470
A;Molecule type: genomic RNA
A;Residues: 1-2230 <TSA>
A;Cross-references: GB:D00924; NID:G225597; PIDN:BAA00766.1; PID:G225598
R;Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H. J. Gen. Virol. 72, 1677-1683, 1991
A;Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and
A;Reference number: JQ1080; MUID:91311420; PMID:1649901
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
submitted to the EMBL Data Library, May 1989
A;Reference number: S04885
A;Accession: S04885
A;Molecule type: genomic RNA
A;Residues: 1750-2164 <BAL1>
A;Cross-references: EMBL:X15461; NID:G61971; PIDN:CAA33490.1; PID:G930268
R;Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhikov
FEBS Lett. 247, 425-428, 1989
A;Title: Variations in genome fragments coding for RNA polymerase in human and simian he
A;Reference number: S03965; MUID:89232168; PMID:2541023
A;Accession: S03965
A;Molecule type: genomic RNA
A;Residues: 1960-2164 <BAL2>
A;Cross-references: EMBL:X15461
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltransferase

RESULT 13

T50995
related to cytoskeleton assembly control protein SLA1 [imported] - Neurospora crassa
N:Alternate names: protein B7F18.140
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T50995
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1119 <SCH>
A:Cross-references: ENBL:AL389891; GSPDB:GN00116; NCSP:B7F18.140
A:Experimental source: BAC clone B7F18; strain OR74A
C:Genetics:
A:Gene: NCSP:B7F18.140
A:Map position: 6
A:Introns: 66/3; 123/2; 495/1

Query Match 49.5%; Score 51; DB 2; Length 1119;

Best Local Similarity 55.6%; Pred. No. 11;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRLKYAQEELNEVLPPP 18

Db 165 QRSVASEDYENVRSP 182

RESULT 14

B82046
proteinase HslVU, ATPase subunit HslU VC2674 [imported] - Vibrio cholerae (strain N16961)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: B82046
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.B.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: B82035; MUID:20406833; PMID:10952301
A:Accession: B82046
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <HEI>
A:Cross-references: GB:AE004333; GB:AE003852; NID:95657266; PIDN:AAF95815.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2674
A:Map position: 1
C:Superfamily: heat shock protein hslU; FtsH/SEC18/CDC48-type ATP-binding domain homolog

Query Match 48.1%; Score 49.5; DB 2; Length 443;

Best Local Similarity 41.7%; Pred. No. 6.8;
Matches 10; Conservative 6; Mismatches 3; Indels 5; Gaps 1;

QY 1 QRLKYAQEELSNE-----VLPPPR 19

Db 117 EKVFRAEELAEERVLALLPPPR 140

RESULT 15

PQ0433
genome polyprotein - human hepatitis A virus (strain PA21) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0433
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0433

A:Molecule type: mRNA

A:Residues: 1-55 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10

Db 46 QRLKYAQEEL 55

RESULT 16

PQ0434
genome polyprotein - human hepatitis A virus (strain KPH) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0434
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0434

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10

Db 47 QRLKYAQEEL 56

RESULT 17

PQ0428
genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0428
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr
A:Reference number: PQ0427; MUID:92300330; PMID:1318940
A:Accession: PQ0428

A:Molecule type: mRNA

A:Residues: 1-56 <ROB>
A:Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polyprotein
C:Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10

Db 47 QRLKYAQEEL 56

RESULT 18

PQ0427
genome polyprotein - human hepatitis A virus (strain EP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: PQ0427

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0427
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
C;Comment: This protein is from the VP1/2A Junction region.
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10
| | | | | | | | | |
Db 47 QRLKYAQEEL 56

RESULT 19
PQ0429
genome polyprotein - human hepatitis A virus (strain PRC16) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0429
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0429
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10
| | | | | | | | | |
Db 47 QRLKYAQEEL 56

RESULT 20
PQ0432
genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0432
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0432
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10
| | | | | | | | | |
Db 47 QRLKYAQEEL 56

RESULT 21
PQ0433
genome polyprotein - human hepatitis A virus (strain 923-1) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0433
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0433
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10
| | | | | | | | | |
Db 47 QRLKYAQEEL 56

PQ0430
genome polyprotein - human hepatitis A virus (strain 923-1) (fragment)
C;Species: human hepatitis A virus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PQ0430
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel J. Gen. Virol. 73, 1365-1377, 1992
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geographical regions
A;Reference number: PQ0427; MUID:92300330; PMID:1318940
A;Accession: PQ0430
A;Molecule type: mRNA
A;Residues: 1-56 <ROB>
A;Note: this protein is from the VP1/2A Junction region
C;Superfamily: hepatitis A virus genome polyprotein
C;Keywords: coat protein; core protein; polyprotein

Query Match 47.6%; Score 49; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEEL 10
| | | | | | | | | |
Db 47 QRLKYAQEEL 56

RESULT 22
S28976
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain - human
A;Alternate names: DNA-directed RNA polymerase B largest chain; DNA-directed RNA polymerase C; Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 18-Jun-1999
C;Accession: S28976; S18986
R;Acker, J.; Wintzerith, M.; Vigneron, M.; Keding, C. J. Mol. Biol. 226, 1295-1299, 1992
A;Title: Primary structure of the second largest subunit of human RNA polymerase II (or A;Reference number: S28976; MUID:92389336; PMID:1518060
A;Accession: S28976
A;Molecule type: mRNA
A;Residues: 1-1174 <ACK>
A;Cross-references: EMBL:X63563; NID:936121; PIDN:CAA45124.1; PID:G36122
C;Superfamily: DNA-directed RNA polymerase 132K polypeptide
C;Keywords: nucleotidyltransferase; transcription

Query Match 47.6%; Score 49; DB 2; Length 1174;
Best Local Similarity 56.2%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNVLP 16
| | | | | | | | | |
Db 334 KRIKYAKEVLOKEMLP 349

RESULT 23
A1969
probable histidine kinase sensor protein - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C;Accession: A1969
R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path A;Reference number: A1969; MUID:99120557; PMID:9923682
A;Accession: A1969
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-442 <ARN>
A;Cross-references: GB:AE001453; GB:AE001439; NID:94154651; PIDN:AAD05720.1; PID:G415465
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0151

Query Match 46.6%; Score 48; DB 2; Length 442;
 Best Local Similarity 75.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 ELSNEVLPPPK 20
 | | | | | | | | | |
 DB 414 EADNEELPPPK 425

RESULT 24
 A27826
 DNA-directed RNA polymerase (EC 2.7.7.6) II second largest chain - fruit fly (Drosophila
 C:Species: Drosophila melanogaster
 C:Date: 19-May-1989 #sequence_revision 23-Nov-1991 #text_change 21-Jul-2000
 C:Accession: A27826; PQ0154
 R:Faikemburg, D.; Dworniczak, B.; Faust, D.M.; Bautz, E.K.F.
 J. Mol. Biol. 195, 929-937, 1987
 A:Title: RNA polymerase II of Drosophila. Relation of its 140,000 Mr subunit to the beta
 A:Reference number: A27826; MUID:88011299; PMID:3116266
 A:Accession: A27826
 A:Molecule type: DNA
 A:Residues: 54-1176 <FAL>
 A:Cross-references: GB:X05709; GB:M29646; NID:G5514651; PIDN:CAA29180.2; PID:G5514652
 R:Sitzler, S.; Oldenburg, I.; Petersen, G.; Bautz, E.K.F.
 Gene 100, 155-162, 1991
 A:Title: Analysis of the promoter region of the housekeeping gene DmRP140 by sequence co
 A:Reference number: PQ0154; MUID:91276237; PMID:1905256
 A:Accession: PQ0154
 A:Molecule type: mRNA
 A:Residues: 1-69 <SIT>
 A:Cross-references: GB:M62972; NID:G157263; PIDN:AAA28476.1; PID:G157267
 A:Experimental source: embryo
 C:Genetics:
 A:Gene: DmRP140
 A:Cross-references: FlyBase:FBgn0003276
 A:Introns: 8/1
 C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
 C:Keywords: DNA binding; nucleotidyltransferase; nucleus; transcription; zinc finger

Query Match 46.6%; Score 48; DB 2; Length 1176;
 Best Local Similarity 56.2%; Pred. No. 36;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QRLKYAQELSNEVLP 16
 | | | | | | | | | |
 DB 335 KRIKYAKEILQKEMLP 350

RESULT 25
 S65068
 DNA-directed RNA polymerase (EC 2.7.7.6) II second largest chain - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 18-Jun-1999
 C:Accession: S65068
 R:Wartlow, D.; Symons, R.H.
 Plant Mol. Biol. 30, 337-342, 1996
 A:Title: Sequence analysis of the second largest subunit of tomato RNA polymerase II.
 A:Reference number: S65068; MUID:96178872; PMID:8616257
 A:Accession: S65068
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1191 <WAP>
 A:Cross-references: EMBL:U28403; NID:G1049067; PIDN:AAC49273.1; PID:G1049068
 C:Genetics:
 A:Gene: RPB2
 C:Superfamily: DNA-directed RNA polymerase 132K polypeptide
 C:Keywords: DNA binding; nucleotidyltransferase; nucleus; transcription

Query Match 46.6%; Score 48; DB 2; Length 1191;
 Best Local Similarity 56.2%; Pred. No. 36;
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 QRLKYAQELSNEVLP 16

Db 338 KRIKYAKEILQKEMLP 353
 | | | | | | | | | |
 RESULT 26
 LLEY
 actin-binding protein ABP1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YCR088w
 C:Species: Saccharomyces cerevisiae
 C:Date: 31-Mar-1991 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
 C:Accession: S19503; S19767; S07608
 R:Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippsen, P.; Schweitzer, B.; Spiegelberg,
 submitted to the Protein Sequence Database, March 1992
 A:Reference number: S19351
 A:Accession: S19503
 A:Molecule type: DNA
 A:Residues: 1-476 <DUS>
 A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R:Frontali, L.; Grisanti, P.
 submitted to the Protein Sequence Database, March 1992
 A:Reference number: S19504
 A:Accession: S19767
 A:Molecule type: DNA
 A:Residues: 14-592 <PRO>
 A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R:Drubin, D.G.; Muhlolland, J.; Zhu, Z.; Botstein, D.
 Nature 343, 288-290, 1990
 A:Title: Homology of a yeast actin-binding protein to signal transduction proteins and m
 A:Reference number: S07608; MUID:90136906; PMID:2405279
 A:Accession: S07608
 A:Molecule type: DNA
 A:Residues: 1-57, 'S', 59-311, 'I', 313-592 <DRU>
 A:Cross-references: EMBL:X51780; NID:G3321; PIDN:CAA36075.1; PID:G3322
 C:Genetics:
 A:Gene: SGD:ABP1; MIPS:YCR088w
 A:Cross-references: SGD:S0000684; MIPS:YCR088w
 A:Map position: 3R
 C:Superfamily: actin-binding protein ABP1; SH3 homology
 C:Keywords: actin binding; duplication
 P:158-180.378-401/Region: duplication
 P:539-588/Domain: SH3 homology <SH3>

Query Match 45.8%; Score 47; DB 1; Length 592;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 QEELSNEVLPPPK 20
 | | | | | | | | | |
 DB 369 EEMENKFPAPPK 382

RESULT 27
 T29757
 protein UNC-89 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 03-Dec-1999
 C:Accession: T29757
 R:Du, Z.; Le, T.T.; Wilson, R.
 submitted to the EMBL Data Library, May 1997
 A:Description: The sequence of C. elegans cosmid C09D1.
 A:Reference number: Z20679
 A:Accession: T29757
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-6642 <DUZ>
 A:Cross-references: EMBL:AF003131; PIDN:AAB54132.1; GSPDB:GN00019; CESP:unc-89
 A:Experimental source: strain Bristol N2; clone C09D1
 C:Genetics:
 A:Gene: CESP:unc-89
 A:Map position: 1
 A:Introns: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1; 61
 /3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6552/3; 6609/1

```

Query Match      45.6%; Score 47; DB 2; Length 6642;
Best Local Similarity 45.0%; Pred. No. 3.7e-02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY      1  QRLKYAEEELSNVEVLPPIRK 20
          | : | | | | | | |
Db      1353  RVSPAEELTPKEIVDSRK 1372

```

RESULT 28
156893
transcription factor AP-4 - human
C.Species: Homo sapiens (man)
C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C.Accession: I56893; A36394
R.Ou, S.H.; Garcia-Martinez, L.F.; Paulsen, E.J.; Gaynor, R.B.
J. Virol. 68, 7188-7199, 1994
A.Title: Role of flanking E box motifs in human immunodeficiency virus type 1 TATA element.
A.Reference number: I56893; PMID:95018629; PMID:7933101

A:Reference number: 158693; PMID:95015629; PMID:19533101
A:Accession: I58693
A:A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-338 <RES>
A:Cross-references: GB:573885; NID:G693848; PIDN:AAB32235.1; PID:G693849
R:Hu, Y. F.; Luescher, B.; Admon, A.; Mermod, N.; Tjian, R.
Genes Dev. 4, 1741-1752, 1990
A:Title: Transcription factor AP-4 contains multiple dimerization domains that regulate
A:Reference number: A36394; PMID:91065520; PMID:2123466
A:Accession: A36394
A:A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 18-338 <HUA>
A:Cross-references: GB:X57435; NID:G37059; PIDN:CAA0683.1; PID:G37060
C:Genetics:
A:Gene: GDB:TFAP4; AP-4
A:Cross-references: GDB:455279; OMIM:600743
A:Map position: 16p13-16p13
C:Keywords: DNA binding; transcription regulation

```

Query Match      44.7%; Score 46; DB 2; Length 338;
Best Local Similarity 38.9%; Pred.No.18;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY      1  QRLKYAQEELSNEVLPPP 18
          ::|::|::|::|::|::|
Dh      210  KRLREOCCLRTOLPPP 227

```

RESULT 29

FWPUB

11S globulin beta subunit precursor - cucurbita cv. Kurokawa Amakuri

N;Alternate names: globulin IV beta subunit

C;Species: Cucurbita cv. Kurokawa Amakuri

A;Variety: cv. Kurokawa Amakuri Nankin

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Sep-2002

C;Accession: S00366; S09480

R;Hayashi, M.; Mori, H.; Nishimura, M.; Akazawa, T.; Hara-Nishimura, I.

Eur. J. Biochem. 172, 627-632, 1988

A;Title: Nucleotide sequence of cloned cDNA coding for pumpkin 11-S globulin beta subunit

A;Reference number: S00366; MUID:88166744; PMID:2450746

A;Accession: S00366

A;Molecule type: mRNA

A;Residues: 1-480 <HAY>

A;Cross-references: EMBL:M36407; NID:g167491; PIDN:AAA33110.1; PID:g167492

A;Experimental source: Cucurbita sp. cv. Kurokawa Amakuri Nankin; cotyledon mRNA

R;Ohmiya, M.; Hara, I.; Matsubara, H.

Plant Cell Physiol. 21, 157-167, 1980

A;Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic and

A;Reference number: S09066

A;Accession: S09480

A;Molecule type: protein

A;Residues: 22-26, B', 28-29, S'; 297-302 <OHM>

A;Experimental source: Cucurbita sp. hybrid Tetsukabuto-Nankin

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:24:26 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-09-171-432A-41

Perfect score: 103

Sequence: 1 QRLKYAQEELSNEVLPPPK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	103	100.0	341	1	POLG_HPAV1	P13672 hepatitis a
2	103	100.0	808	1	POLG_HPAV1	Q02381 hepatitis a
3	103	100.0	852	1	POLG_HPAV1	P05442 hepatitis a
4	103	100.0	2226	1	POLG_HPAV2	P26580 hepatitis a
5	103	100.0	2226	1	POLG_HPAV2	P26581 hepatitis a
6	103	100.0	2226	1	POLG_HPAV8	P26581 hepatitis a
7	103	100.0	2227	1	POLG_HPAV8	P08617 hepatitis a
8	103	100.0	2227	1	POLG_HPAV1	P05441 hepatitis a
9	103	100.0	2227	1	POLG_HPAV1	P13901 hepatitis a
10	97	94.2	2230	1	POLG_HPAV5	P14553 simian hepa
11	94	91.3	839	1	POLG_HPAV1	P31788 simian hepa
12	52	50.5	1193	1	RPB2_CAEEL	Q10578 caenorhabdi
13	49.5	48.1	443	1	HSJU_VIBCH	Q9knq7 vibrio chol
14	49.5	48.1	443	1	HSJU_VIBVU	Q8dcp4 vibrio vuln
15	49	47.6	1174	1	RPB2_HUMAN	P30876 homo sapien
16	48	46.6	636	1	NAF1_HUMAN	Q15025 homo sapien
17	48	46.6	1176	1	RPB2_DROME	P08266 drosophila
18	48	46.6	1191	1	RPB2_LYCES	Q42877 lycopersico
19	47	45.6	592	1	ABP1_YEAST	P15891 saccharomyc
20	47	45.6	6832	1	UN99_CAEEL	Q01761 caenorhabdi
21	46	44.7	338	1	TAP4_HUMAN	Q01664 homo sapien
22	46	44.7	480	1	11S8_CUCMA	P13744 cucurbita m
23	45	43.7	733	1	HEXA_BLADI	Q17127 blaberius di
24	45	43.7	1188	1	RPB2_ARATH	P38420 arabidopsis
25	44	42.7	303	1	PMG3_YEAST	Q12326 saccharomyc
26	44	42.7	443	1	HSJU_YERPE	Q82j55 yersinia pe
27	44	42.7	510	1	GPW1_ANTSP	Q06464 antithamio
28	44	42.7	914	1	ORC1_YEAST	P54784 saccharomyc
29	43.5	42.2	3678	1	DMD_MOUSE	P11531 mus musculu
30	43.5	42.2	3680	1	DMD_CANFA	Q97592 canis fami
31	43.5	42.2	3685	1	DMD_HUMAN	P11532 homo sapien
32	43	41.7	120	1	YLOU_BACSU	Q34318 bacillus su
33	43	41.7	268	1	YHC6_YEAST	P38740 saccharomyc

RESULT 1

ALIGNMENTS

34	43	41.7	574	1	HEMA_INBMD	P03461 influenza b
35	43	41.7	576	1	HEMA_INBUS	P09766 influenza b
36	43	41.7	578	1	HEMA_INBME	P09765 influenza b
37	43	41.7	578	1	HEMA_INBVI	P09767 influenza b
38	43	41.7	583	1	HEMA_INBEN	P10757 influenza b
39	43	41.7	583	1	HEMA_INBOR	P03464 influenza b
40	43	41.7	583	1	HEMA_INBSI	P03463 influenza b
41	43	41.7	585	1	HEMA_INBSE	P17504 influenza b
42	43	41.7	585	1	HEMA_INBEV	P22092 influenza b
43	43	41.7	1146	1	AS10_YEAST	P48361 saccharomyc
44	43	41.7	1210	1	RPB2_SCHPO	Q02061 schizosacch
45	43	41.7	2696	1	NSD1_HUMAN	Q06173 homo sapien
46	42.5	41.3	2278	1	FABI_YEAST	P34756 saccharomyc
47	42	40.8	214	1	DEVS_MTXXA	Q07766 myxococcus
48	42	40.8	276	1	YJIC_ECOLI	P39374 escherichia
49	42	40.8	445	1	ML64_HUMAN	Q14849 homo sapien
50	42	40.8	554	1	PEX2_PODAN	P51021 podospora a
51	42	40.8	1005	1	DPOL_VARV	P33793 variola vir
52	42	40.8	1006	1	DPOL_VACCC	P20509 vaccinia vi
53	42	40.8	1006	1	DPOL_VACCV	P06856 vaccinia vi
54	42	40.8	1066	1	AOSL_PLEHO	Q16025 plexaura ho
55	42	40.8	1270	1	DXH9_HUMAN	Q08211 homo sapien
56	42	40.8	1773	1	DIP2_DROME	Q9W089 drosophila
57	41.5	40.3	348	1	OMH1_PASMU	Q9Cm99 pasteurella
58	41.5	40.3	498	1	PM2_DROME	P13369 drosophila
59	41	39.8	181	1	ATPF_STRLI	P50013 streptomyce
60	41	39.8	258	1	RT15_MOUSE	Q9dc71 mus musculu
61	41	39.8	321	1	PE28_ARATH	Q9S677 arabidopsis
62	41	39.8	349	1	F161_RHIME	Q9S674 rhizobium m
63	41	39.8	349	1	F162_RHIME	P56886 rhizobium m
64	41	39.8	451	1	CSSS_BACSU	Q32193 bacillus su
65	41	39.8	460	1	ANT3_HUMAN	Q9Y5c1 homo sapien
66	41	39.8	500	1	CISA_BACSU	P17867 bacillus su
67	41	39.8	512	1	FLOH_MOUSE	P41438 mus musculu
68	41	39.8	519	1	TRPE_SERMA	P00897 serratia ma
69	41	39.8	758	1	YP58_YEAST	Q99299 saccharomyc
70	41	39.8	789	1	ATX1_RAT	Q63540 rattus norv
71	41	39.8	816	1	ATX1_HUMAN	P54253 homo sapien
72	41	39.8	1719	1	PRD2_HUMAN	Q13029 homo sapien
73	41	39.8	1966	1	CCAF_HUMAN	O60840 homo sapien
74	41	39.8	1985	1	CCAF_MOUSE	Q9J1s7 mus musculu
75	40.5	39.3	327	1	HUNE_MANSE	Q25514 manduca sex
76	40	38.8	319	1	DBP8_YEAST	P38637 saccharomyc
77	40	38.8	443	1	Y314_MYCGE	Q49415 mycoplasma
78	40	38.8	584	1	HCM1_YEAST	P25364 saccharomyc
79	40	38.8	624	1	ARLY_VIBVU	Q8dcm9 vibrio vuln
80	40	38.8	681	1	COG2_CAEEL	Q21444 caenorhabdi
81	40	38.8	802	1	OPH1_MOUSE	Q99731 mus musculu
82	40	38.8	853	1	LOX4_SOYBN	P38417 glycine max
83	40	38.8	935	1	MYCD_MOUSE	Q8vims mus musculu
84	40	38.8	938	1	MYCD_RAT	Q8vsi7 rattus norv
85	40	38.8	1097	1	CCT_DROME	O96433 drosophila
86	40	38.8	1197	1	SMC2_HUMAN	O95347 homo sapien
87	40	38.8	1229	1	SIP3_YEAST	P38717 saccharomyc
88	40	38.8	1380	1	DXK3_MOUSE	O70133 mus musculu
89	40	38.8	1398	1	TRC3_MOUSE	O90900 m nuclear r
90	40	38.8	2032	1	NGPG_CAEEL	Q93971 caenorhabdi
91	40	38.8	2390	1	SPCP_HUMAN	O15020 homo sapien
92	40	38.8	2476	1	ATRX_MOUSE	O61687 mus musculu
93	39.5	38.3	482	1	YSR2_CAEEL	Q09950 caenorhabdi
94	39.5	38.3	530	1	GME2_HUMAN	Q9ukd1 homo sapien
95	39.5	38.3	702	1	CRN_DROME	P17886 drosophila
96	39.5	38.3	867	1	NIA_ASPNG	P36858 aspergillus
97	39.5	38.3	908	1	ATME_SALTY	P22036 salmonella
98	39.5	38.3	931	1	MCMA_SCHPO	P29458 schizosacch
99	39.5	38.3	1589	1	PHP_DROME	P19769 drosophila
100	39.5	38.3	2241	1	TEGU_HCMVA	P16785 human cytom

```
POLG_HPAV1
ID POLG_HPAV1 STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP3 (1C); Coat protein VP1 (1D); Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LCDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Aronov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
FT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14666; CAA32794.1; -.
CC PIR; S04137.
CC InterPro; IPR008975; Viral_cap_coat.
CC Polyprotein; Coat protein; Core protein.
CC NON_TER 1 1 COAT PROTEIN VP3.
CC CHAIN <1 1 COAT PROTEIN VP1.
CC CHAIN 2 340 COAT PROTEIN P2A.
CC CHAIN 341 >341 CORE PROTEIN P2A.
CC NON_TER 341 341
CC SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;
Query Match 100.0%; Score 103; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 1.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 314 QRLKYAQEELSNEVLPPPRK 333
RESULT 2
POLG_HPAVG
ID POLG_HPAVG STANDARD; PRT; 808 AA.
AC Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain CA76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=31706;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
```

```
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M66695; AAA45477.1; -.
CC InterPro; IPR008975; Viral_cap_coat.
CC Polyprotein; Coat protein; Core protein.
CC NON_TER 1 1 COAT PROTEIN VP4 (P1A).
CC CHAIN <1 2 COAT PROTEIN VP2 (P1B).
CC CHAIN 3 223 COAT PROTEIN VP3 (P1C).
CC CHAIN 224 470 COAT PROTEIN VP1 (P1D).
CC CHAIN 471 770 COAT PROTEIN P2A.
CC CHAIN 771 >808 CORE PROTEIN P2A.
CC NON_TER 808 808
CC SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;
Query Match 100.0%; Score 103; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 789 QRLKYAQEELSNEVLPPPRK 808
RESULT 3
POLG_HPAVC
ID POLG_HPAVC STANDARD; PRT; 852 AA.
AC P06442; Q83741; Q83742;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12097;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85185648; PubMed=2985793;
RA Lineneyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M10033; AAA45470.1; -.
CC PIR; A03904; GNHYHA.
CC InterPro; IPR008975; Viral_cap_coat.
CC Polyprotein; Coat protein; Core protein.
CC CHAIN 1 23 COAT PROTEIN VP4 (P1A).
CC CHAIN 24 245 COAT PROTEIN VP2 (P1B).
CC CHAIN 246 491 COAT PROTEIN VP3 (P1C).
```

```

FT CHAIN 1088 1422 CORE PROTEIN P2C.
CH14 CHAIN 1423 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D..
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 103; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLYAQEELSNEVLPPPRK 20
|||||
810 ORLYAQEELSNEVLPPPRK 829

```

```

FT CHAIN 492 836 COAT PROTEIN VP1 (P1D) .
FT CHAIN 837 >852 CORE PROTEIN P2A.
FT FT 852 852
FT NON TER 852
FT SQ SEQUENCE 852 AA; 95563 NW; 793DED0AD532820E CRC64;

Query Match 100.0%; Score 103; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 QELKYAQBELSNEVLPPRK 20
|||||
Db 810 QRLKYAQBELSNEVLPPRK 829

```

RESULT 5	
POLG_HPAV4	
ID	POLG_HPAV4
STANDARD;	
PRT;	2226 AA.
AC	P26581;
DT	01-AUG-1992 (Rel. 23, Created)
DDT	01-AUG-1992 (Rel. 23, Last sequence update)
DE	28-FEB-2003 (Rel. 41, Last annotation update)
DE	Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS	Hepatitis A virus (strain 43C).
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.
NCBI	maxID=12095

POLG_HPAV2	STANDARD;	PRT: 2226 AA.
P26580;		
01-AUG-1992	(Rel. 23, Created)	
01-AUG-1992	(Rel. 23, Last sequence update)	
28-FEB-2003	(Rel. 41, Last annotation update)	
Genome polyprotein	[Contains: Coat proteins VP1 TO VP4; Core proteins	
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase		
P3D (EC 2.7.7.48)]		
Hepatitis A virus (strain 24a).		
Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
Hepatovirus.		
NCBI_TaxID=12094;		
[1]		

SEQUENCE FROM N.A.
MEDLINE=91162759; PubMed=1705995;
Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
Cromeans T., Jansen R.W.;
"Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination.";
J. Virol. 65:2056-2065(1991).
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
(RNA) (N).
-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RA "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination";
RT J. Virol. 65:2056-2065 (1991).
RL -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC -----

CC This Swiss-ProJ entry is copyrighted, it is produced and distributed under the terms of the CC-BY-NC-ND 4.0 International license.
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; M59809; AAA45469.1; --
 CC MEROPS; C03.005; --
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR000605; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; MS9810; AAA45468.1; --
DR MEROPS; C03.005; --
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_tryp.sin.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_P5v1r.
DR

DR	InterPro; LPX008975; viral_cap_cocat.
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR	Pfam; PF00910; RNA_helicase; 1.
DR	PRINTS; PR00918; CALICVIRUSNS.
DR	Polypeptide; Coat protein; Core protein; Transferase;
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT	CHAIN 1 23
FT	CHAIN 24 245
FT	CHAIN 246 491
FT	CHAIN 492 794

DR	PFam: PF00910; RNA helicase: 1.
DR	PRINTS; PRO0918; CALICVIRUSNS.
KW	Polyprotein; Coat protein; Core protein; Transferase;
KW	RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT	CHAIN 1 23
FT	CHAIN 24 245
FT	CHAIN 246 491
FT	CHAIN 492 794
FT	CHAIN 795 900
FT	CHAIN 901 1087
FT	CORE PROTEIN VP2 (PIB).
FT	CORE PROTEIN VP3 (PIC).
FT	CORE PROTEIN VP1 (PID).
FT	CORE PROTEIN P2A.
FT	CORE PROTEIN P2B.

```

FT CHAIN      24    245          COAT PROTEIN VP2 (P1B).
FT CHAIN     246    491          COAT PROTEIN VP3 (P1C).
FT CHAIN     492    794          COAT PROTEIN VP1 (PID).
FT CHAIN     795    900          CORE PROTEIN P2A.
FT CHAIN     901   1087          CORE PROTEIN P2B.
FT CHAIN    1088   1422          CORE PROTEIN P2C.
FT CHAIN    1423   1915          PROBABLE PROTEIN P3A.
FT CHAIN    1496   1518          PROBABLE PROTEIN P3B.
FT CHAIN    1519   1737          PROBABLE PROTEIN P3C.
FT CHAIN    1738   2226          RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE   2226 AA; 251292 MW; 24964A63396CD6B CRC64;

Query Match      100.0%; Score 103; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ORLKYAQEELSGNEVLPPPRK 20
DB      810 QRLKYAQEELSGNEVLPPPRK 829
        |||||
RESULT 7
POLG.HPAHVH STANDARD; PRT; 2227 AA.
AC F08617; P06443; Q81082;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain HM-175).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI TaxID=12098;
[1]_TaxID=12098;
RN SEQUENCE FROM N.A.
RP STRAIN=Wild type;
RC MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
RA Baroudy B.M.;
RT "Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.";
RL J. Virol. 61:50-59 (1987).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Attenuated;
RC MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,
RA Purcell R.H.;
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501 (1987).
[3]
RN SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RP MEDLINE=85166289; PubMed=2984684;
RX Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
RA Purcell R.H., Feinstone S.M.;
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147 (1985).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.

```

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR	EMBL; M44114; AAA45475.1; --	
DR	EMBL; M44707; AAA45465.1; --	
DR	EMBL; M44707; AAA45466.1; ALT_INIT.	
DR	EMBL; M16632; AAA45471.1; --	
DR	PIR; A03905; A03905.	
DR	FIR; A25981; GNNYH.	
DR	PIR; A94149; GNNYK.	
DR	PDB; 1HAV; 23-DEC-96.	
DR	MEROBS; C03.005; --	
DR	InterPro; IPR004004; Calici_pol_hel.	
DR	InterPro; IPR009003; Cys_Ser_trypsin.	
DR	InterPro; IPR000605; RNA_helicase.	
DR	InterPro; IPR007095; RNA_pol_DS_PS.	
DR	InterPro; IPR001205; RNA_pol_P3D.	
DR	InterPro; IPR007094; RNA_pol_Psvir.	
DR	InterPro; IPR008975; Viral_cap_coat.	
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.	
DR	Pfam; PF00910; RNA_helicase; 1.	
DR	PRINTS; PR0918; CALICIVIRUS.	
KW	Polypeptide; Coat protein; Core protein; Transferase;	
KW	RNA-directed RNA polymerase; Hydrolase; Triol protease; 3D-structure.	
FT	CHAIN 1 23	
FT	CHAIN 24 245	COAT PROTEIN VP4 (PIA).
FT	CHAIN 246 491	COAT PROTEIN VP2 (PIB).
FT	CHAIN 492 836	COAT PROTEIN VP3 (PIC).
FT	CHAIN 837 980	COAT PROTEIN VP1 (PID).
FT	CHAIN 981 1087	CORE PROTEIN P2A.
FT	CHAIN 1088 1422	CORE PROTEIN P2B.
FT	CHAIN 1423 1496	CORE PROTEIN P2C.
FT	CHAIN 1497 1519	PROBABLE PROTEIN P3A.
FT	CHAIN 1520 1738	PROBABLE PROTEIN P3B.
FT	CHAIN 1739 2227	PROBABLE PROTEIN P3C.
FT	VARIANT 77 77	RNA-DIRECTED POLYMERASE 3D.
FT	VARIANT 764 764	X -> R (IN ATTENUATED STRAIN).
FT	VARIANT 821 821	E -> V (IN ATTENUATED STRAIN).
FT	VARIANT 1052 1052	N -> S (IN ATTENUATED STRAIN).
FT	VARIANT 1062 1062	A -> V (IN ATTENUATED STRAIN).
FT	VARIANT 1118 1118	G -> A (IN ATTENUATED STRAIN).
FT	VARIANT 1151 1151	K -> M (IN ATTENUATED STRAIN).
FT	VARIANT 1163 1163	E -> K (IN ATTENUATED STRAIN).
FT	VARIANT 1277 1277	F -> S (IN ATTENUATED STRAIN).
FT	VARIANT 1277 1277	V -> I (IN ATTENUATED STRAIN).
FT	VARIANT 1500 1500	H -> Y (IN ATTENUATED STRAIN).
FT	VARIANT 1805 1805	D -> N (IN ATTENUATED STRAIN).
FT	VARIANT 1930 1930	S -> T (IN ATTENUATED STRAIN).
SQ	SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8B740A6 CRC64;	

```

Query Match      100.0%; Score 103; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 QRLKYAQEELSNEVLPPPRK 20 QY

Db 810 QRLKYAQEELSNEVLPPRK 829

RESULT 8

RESOL 8
POLG HPAYL

POLG_HEAVY
ID_POLG

AC P0644:

DT 01-JAN

DT 01-JAN

DT 28-FEB

DE
Genome

DE P2A TO

P3D (EC 2.7.7.48)].
Hepatitis A virus (strain LA).
Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=12099;
[1]
SEQUENCE FROM N.A.
MEDLINE=85190549; PubMed=2986127;
Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
Merryweather J., van Nest G., Dina D.;
"Primary structure and gene organization of human hepatitis A virus.";
Proc. Natl. Acad. Sci. U.S.A. 82:2629-2631(1985).
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA}[N].
-!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 each of which is composed of one copy each of proteins VP1, VP2,
 VP3, and VP4.
-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL; K02990; AAA45472.1; -.
PIR; A03903; GNNVHR.
MEROFS; C03.005; -.
InterPro; IPR004004; Calici_pol_hel.
InterPro; IPR009003; Cys_Ser_trypsin.
InterPro; IPR000605; RNA_helicase.
InterPro; IPR007895; RNA_pol_DS_PS.
InterPro; IPR001205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_PSVir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00680; RNA_dep_RNA_pol_1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICIVIRUSNS.
Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease.
CHAIN 1 23
CHAIN 24 245
CHAIN 246 491
CHAIN 432 836
CHAIN 432 836
CHAIN 837 981
CHAIN 981 1076
CHAIN 1077 1422
CHAIN 1423 1484
CHAIN 1485 1507
CHAIN 1508 1678
CHAIN 1679 2227
2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;
SEQUENCE

Query Match	100.0%;	Score 103;	DB 1;	Length 2227;
Best Local Similarity	100.0%;	Pred. NO. 8.8e-08;		
Matches 20:	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qv 1 ORLKVAOEEELSNEVLPPRK 20

810 QRI.KYACERTSNEVLPPPRK 829

DECLASS. AUTHORITY: 25 C.F.R. 171.22

RESULT 9
POLIG HPAYM

POLG_HPAVM
TD_POIG

AC P1390

AC
ACAC DT
01-JAN
081030

DT 01-JAN

DT 28-FEB

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ORLKYAEELSNEVLPPPRK 20
 DB 814 ORLKYAEELSNEVLPPPRK 833

RESULT 11
 POLG_HPAVT STANDARD; PRT; 839 AA.
 AC P31788;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 DE P2A] (Fragment).
 OS Simian hepatitis A virus (strain CY-145).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus
 OC NCBI_TaxID=31707;
 [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=91311421; PubMed=1649902;
 RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
 RT "Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus macaques (Macaca fascicularis).";
 RL J. Gen. Virol. 72:1685-1689 (1991).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
 CC -1- PRT: Specific enzymatic cleavages in vivo yield mature proteins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M59286; AAA45473.1; --
 DR PIR; JQ1180; GNNYS2.
 DR InterPro; IPR008975; Viral_cap_coat.
 DR Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
 FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
 FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
 FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
 FT CHAIN ? >839 CORE PROTEIN P2A.
 FT NON_TER 839 839
 SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;
 Query Match 91.3%; Score 94; DB 1; Length 839;
 Best Local Similarity 85.0%; Pred. No. 7.8e-07;
 Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ORLKYAEELSNEVLPPPRK 20
 DB 809 ORFKYAEELSNEVLPPPRK 828

RESULT 12
 RB2_CABEL STANDARD; PRT; 1193 AA.
 AC Q10578;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE DNA-directed RNA polymerase II second largest subunit (EC 2.7.7.6)
 DE (RNA polymerase I subunit 2).
 DE C26E6.4.
 GN C26E6.4.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

Rhabditidae; Peloderinae; Caenorhabditis.

OC NCBI_TaxID=6239;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=Bristol N2;
 RC Fulton L.;
 RA Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RN SEQUENCE OF 200-1058 FROM N.A.
 RX MEDLINE=95041334; PubMed=7953533;
 RA Sidow A., Thomas W.K.;
 RT "A molecular evolutionary framework for eukaryotic model organisms.";
 RL Curr. Biol. 4:596-603 (1994).
 CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates.
 CC
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}(N).
 CC
 CC -1- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14 different polypeptides. This subunit is the second largest component of RNA polymerase II.
 CC
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC
 CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are found in eukaryotic nuclei: polymerase I for the ribosomal RNA precursor, polymerase II for the mRNA precursor, and polymerase III for 5S and tRNA genes.
 CC
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U13975; AAA21158.1; --
 DR EMBL; U10333; AAA50224.1; --
 DR PIR; B88445; B88445.
 DR PIR; T43701; T43701.
 DR WormPep; C26E6.4; CE01162.
 DR InterPro; IPR007121; RNA_pol_B.
 DR InterPro; IPR007644; RNA_pol_Rpb2_1.
 DR InterPro; IPR007642; RNA_pol_Rpb2_2.
 DR InterPro; IPR007645; RNA_pol_Rpb2_3.
 DR InterPro; IPR007646; RNA_pol_Rpb2_4.
 DR InterPro; IPR007647; RNA_pol_Rpb2_5.
 DR InterPro; IPR007120; RNA_pol_Rpb2_6.
 DR InterPro; IPR007641; RNA_pol_Rpb2_7.
 DR Pfam; PF04563; RNA_pol_Rpb2_1; 1.
 DR Pfam; PF04561; RNA_pol_Rpb2_2; 1.
 DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
 DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
 DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
 DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
 DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
 DR PROSITE; PS01166; RNA_POL_BETA; 1.
 KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
 KW Zinc-finger; Nuclear protein.
 KW ZN_FING 1125 1146
 SQ SEQUENCE 1193 AA; 134904 MW; B8A85E74E9C7EBE CRC64;
 Query Match 50.5%; Score 52; DB 1; Length 1193;
 Best Local Similarity 62.5%; Pred. No. 4.7;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 ORLKYAEELSNEVLPP 16
 DB 340 ORIKYAREILQXELLP 355

RESULT 13
 HSLU_VIBCH

ID HSLU_VIBCH STANDARD; PRT; 443 AA.
 AC O9KN07;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP-dependent hsl protease ATP-binding subunit hslu.
 GN HSLU OR VC2674.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Emolae M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Ueberback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RA "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RT Nature 406:477-483 (2000).
 RL -1- FUNCTION: Chaparone subunit of a proteasome-like degradation
 CC complex (By similarity).
 CC -1- SUBUNIT: A double ring-shaped homohexamers of hslu is capped on
 CC each side by a ring-shaped hslu homohexamers (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the clpX chaparone family. Hslu subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE004333; AAF95815.1; -
 DR PIR: E82046; E82046.
 DR HSP: P32168; IDO2.
 DR TIGR: VC2674; -
 DR HAMAP: MF 00249; -; 1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003959; AAA_ATPase_cent.
 DR InterPro: IPR004491; Hsp_HsluV.
 DR Pfam: PF00004; AAA; 1.
 DR SMART: SM00382; AAA; 1.
 DR TIGRFAMs: TIGR00390; hslu; 1.
 DR TrEMBL: IPR003593; AAA_ATPase.
 KW Chaparone, ATP-binding; Complete proteome.
 FT NP_BIND 57 64 ATP (POTENTIAL).
 SQ SEQUENCE 443 AA; 49900 MW; DA13E82FA6A38F CRC64;
 Query Match 48.1%; Score 49.5; DB 1; Length 443;
 Best Local Similarity 41.7%; Pred. No. 3.9;
 Matches 10; Conservative 6; Mismatches 3; Indels 5; Gaps 1;
 Qy 1 ORLKYAQBELSNE-----VLPPPP 19
 Db 117 EKVFRAEELAEERVLDAALLPPPP 140
 RESULT 14
 HSLU_VIBU
 ID HSLU_VIBU STANDARD; PRT; 443 AA.
 AC Q8DCP4;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ATP-dependent hsl protease ATP-binding subunit hslu.
 GN HSLU OR VV11355.

OS Vibrio vulnificus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=672;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CMCP6;
 RA Choe H.E.;
 RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
 RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Chaparone subunit of a proteasome-like degradation
 CC complex (By similarity).
 CC -1- SUBUNIT: A double ring-shaped homohexamers of hslu is capped on
 CC each side by a ring-shaped hslu homohexamers (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the clpX chaparone family. Hslu subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE016801; AAO09806.1; -
 DR HAMAP: MF 00249; -; 1.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003959; AAA_ATPase_cent.
 DR InterPro: IPR004491; Hsp_HsluV.
 DR Pfam: PF00004; AAA; 1.
 DR SMART: SM00382; AAA; 1.
 DR TIGRFAMs: TIGR00390; hslu; 1.
 DR TrEMBL: IPR003593; AAA_ATPase.
 KW Chaparone, ATP-binding; Complete proteome.
 FT NP_BIND 57 64 ATP (POTENTIAL).
 SQ SEQUENCE 443 AA; 49993 MW; 634F55B3AC85F23D CRC64;
 Query Match 48.1%; Score 49.5; DB 1; Length 443;
 Best Local Similarity 41.7%; Pred. No. 3.9;
 Matches 10; Conservative 6; Mismatches 3; Indels 5; Gaps 1;
 Qy 1 ORLKYAQBELSNE-----VLPPPP 19
 Db 117 EKVFRAEELAEERVLDAALLPPPP 140
 RESULT 15
 RPB2_HUMAN
 ID RPB2_HUMAN STANDARD; PRT; 1174 AA.
 AC P30876;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE DNA-directed RNA polymerase II 140 kDa polypeptide (EC 2.7.7.6)
 DE (RNA polymerase II subunit 2) (RPS2).
 GN POLR2B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92389336; PubMed=1518060;
 RA Acker J., Wintzerith M., Vigneron M., Keding C.;
 RT "Primary structure of the second largest subunit of human RNA
 RT polymerase II (or B).";
 RL J. Mol. Biol. 226:1295-1299 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
QN zinc-finger; Nuclear protein.
FT ZN FING 1119 1140 C4-TYPE (POTENTIAL).
SQ SEQUENCE 1174 AA; 133896 MW; 32BEDF7F95E4DE10 CRC64;

Query Match 47.6%; Score 49; DB 1; Length 1174;
Best Local Similarity 56.2%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORLKVAQEELSNVLP 16
| | | | | | | | | | | | | | | | | |
DB 334 KRKIYAKEVLQKEMLP 349

RESULT 16
NAFI_HUMAN
ID NAFI_HUMAN STANDARD; PRT; 636 AA.
AC Q15025; OY6008; Q96EL9; Q99833; Q9HJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Nef-associated factor 1 (Naf1) (HIV-1 Nef interacting protein)
DE (Virion-associated nuclear shuttling protein) (VAN) (hVAN) (TNFAIP3
DE interacting protein 1).
DE TNF1 OR NAF1 OR KIAA0113.
GN Homo sapiens (Human). OS
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RS SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP TISSUE=PerIPHERAL blood;
RC MEDLINE=99120485; PubMed=9923610;
RX Fukushi M., Dixon J., Kimura T., Tsurutani N., Dixon M.J.,
RA Yamamoto N.;
RT "Identification and cloning of a novel cellular protein Naf1, Nef-
RT associated factor 1, that increases cell surface CD4 expression.";
RT FEBS Lett. 442:83-88(1999).
RN [2]
RS SEQUENCE FROM N.A. (ISOFORM 1).
RP TISSUE=Leukocyte;
RC MEDLINE=20541981; PubMed=11090181;
RX Gupta K., Ott D., Hope T.J., Siliciano R.F., Boeke J.D.;
RA "A human nuclear shuttling protein that interacts with human
RT immunodeficiency virus type 1 matrix is packaged into virions.";
RL J. Virol. 74:11811-11824(2000).
RN [3]
RS SEQUENCE FROM N.A. (ISOFORM 1).
RP TISSUE=Lung;
RC MEDLINE=22388257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zeeberg B., Bustov K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettenman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RS SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).
RP

CC TISSUE=Craniofacial;
RX MEDLINE=96276047; PubMed=8681136;
RA Loftus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;
RT "Transcriptional map of the Treacher Collins candidate gene region.";
RL Genome Res. 6:26-34(1996).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RA Ohara O., Nagase T., Kikuno R., Nomura N.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE OF 341-636 FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow;
RX MEDLINE=95308325; PubMed=7798527;
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,
RA Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. III.
RT The coding sequences of 40 new genes (K1AA0081-K1AA0120) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:37-43(1995).
RN [7]
RP SEQUENCE OF 94-412 FROM N.A.
RA Fukushi M., Kimura T., Yamamoto N.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and
CC inhibits TNF-induced NF-kappa-B-dependent gene expression by
CC interfering with an RIP- or TRAF2-mediated transactivation signal
CC (By similarity). Increases cell surface CD4(T4) antigen
CC expression. Interacts with HIV-1 matrix protein and is packaged
CC into virions and overexpression can inhibit viral replication. May
CC regulate matrix nuclear localization, both nuclear import of PIC
CC (preintegration complex) and export of GAG polyprotein and viral
CC genomic RNA during virion production.
CC -1- SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with
CC HIV-1 matrix protein.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus
CC and cytoplasm in a CRM1-dependent manner.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=Alpha;
CC IsoId=Q15025-1; Sequences=Displayed;
CC Name=2; Synonyms=Beta;
CC IsoId=Q15025-2; Sequences=VSP 003913;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Ubiquitous. Strongly expressed in peripheral
CC blood lymphocytes, spleen and skeletal muscle, and is weakly
CC expressed in the brain.
CC -1- CAUTION: Ref.7 sequence differs from that shown due to frameshifts
CC in positions 152 and 154.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ011895; CAA09855.1; -
DR EMBL: AJ011896; CAA09856.1; -
DR EMBL: AJ012155; AAG42154.1; -
DR EMBL: BC012133; AAH12133.1; -
DR EMBL: BC014008; AAH14008.1; -
DR EMBL: U39403; AAC99999.1; -
DR EMBL: D30755; BAA06416.2; -
DR EMBL: U03844; AAB41438.1; ALT_FRAME.
DR GenBank: HGNC:16903; TNIP1.
DR MIM: 607714; -
DR GO: GO:0005622; C:intracellular; TAS.
DR GO: GO:000515; F:protein binding; TAS.
DR GO: GO:0003101; P:glycoprotein biosynthesis; IDA.
DR GO: GO:0045071; P:negative regulation of viral genome replication; TAS.
DR GO: GO:0045071; P:negative regulation of viral genome replication; TAS.
KW Coiled coil; Nuclear protein; Alternative splicing.
FT DOMAIN 20 73 COILED COIL (POTENTIAL).

FT DOMAIN 196 258 COILED COIL (POTENTIAL).
FT DOMAIN 294 535 COILED COIL (POTENTIAL).
FT DOMAIN 94 412 INTERACTS WITH NEF.
FT DOMAIN 524 530 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 539 636 PRO-RICH.
FT VARSPLIC 627 636 SPKNDREGQP -> PADLRPRN (in isoform 2).
FT CONFLICT 148 148 G -> D (IN REF. 3; AAH12133).
FT CONFLICT 299 299 A -> P (IN REF. 2).
SQ SEQUENCE 636 AA; 71864 MW; D81B96BEAD50D871 CRC64;
Query Match 46.6%; Score 48; DB 1; Length 636;
Best Local Similarity 57.9%; Pred.No. 9.9;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;
Oy 2 RLKYAQEEL--SNEVLPPP 18
||: ||| ||: |||
Db 58 RLKQAEELVKNELPPP 76

RESULT 17

ID RPB2_DROME STANDARD; PRT; 1176 AA.
AC P08266; Q04155; Q95027; Q9VFM7;
DT 01-AUG-1988 (Rel. 08, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA-directed RNA polymerase II 140 kDa polypeptide (EC 2.7.7.6)
DE (RNA polymerase II subunit 2).
GN RPII140 OR CG3180.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=88011299; PubMed=3116266;
RA Falkenburg D., Dworkiczak B., Faust D.M., Bautz E.K.F.;
RT "RNA polymerase II of Drosophila. Relation of its 140,000 Mr subunit
RT to the beta subunit of Escherichia coli RNA polymerase.";
RL J. Mol. Biol. 195:929-937(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Roarkins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,

```

InterPro; IPR007641; RNA_pol_Rpb2_7.
Pfam; PF04563; RNA_pol_Rpb2_1; 1.
Pfam; PF04561; RNA_pol_Rpb2_2; 1.
Pfam; PF04565; RNA_pol_Rpb2_3; 1.
Pfam; PF04566; RNA_pol_Rpb2_4; 1.
Pfam; PF04567; RNA_pol_Rpb2_5; 1.
Pfam; PF00562; RNA_pol_Rpb2_6; 1.
Pfam; PF04560; RNA_pol_Rpb2_7; 1.
PROSITE; PS01166; RNA_POL_BETA; 1.
Transferrase; DNA-directed RNA polymerase; Transcription; Zinc.
Zinc-finger; Metal-binding; Nuclear protein.
ZN_FING 1121 1142 C4-TYPE (POTENTIAL).
CONFLICT 72 72 A -> R (IN REF. 1).
CONFLICT 666 667 ID -> MY (IN REF. 1).
SEQUENCE 1176 AA; 134042 MW; 224821B335BED7F0 CRC64;

Query Match 46.6%; Score 48; DB 1; Length 1176;
Best Local Similarity 56.2%; Pred.No.20;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 ORLYAQEELSNVLP 16
DB 335 KRIKYAEILQKEMLP 350

RESULT 18
RBP2_LYC8S STANDARD; PRT; 1191 AA.
AC Q42877;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase II 135 kDa polypeptide (EC 2.7.7.6)
DE (RNA polymerase II subunit 2).
DE RPB2.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OC NCBI_TaxID=4081;
RN [1]_TaxID=4081;
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Grosse lisse;
RC MEDLINE=96178872; PubMed=8616257;
RA Warrilow D., Symons R.H.;
RT "Sequence analysis of the second largest subunit of tomato RNA
RT polymerase II."
RL Plant Mol. Biol. 30:337-342 (1996).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
CC different polypeptides. This subunit is the second largest
CC component of RNA polymerase II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for 5S and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U28403; AAC49273.1; --
CC FIR; S65068; S65068.

```


lattice assembly begins with positional cues laid down in the basement membrane and muscle cell membranes. UNC-89 responds to these signals, localizes, and then participates in assembling an M-line.

-!- TISSUE SPECIFICITY: Localizes to the middle of A-bands.

-!- SIMILARITY: Contains 1 DBL-homology (DB) domain.

-!- SIMILARITY: Contains 1 fibronectin type III domain.

-!- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.

-!- SIMILARITY: Contains 1 PH domain.

-!- SIMILARITY: Contains 5 RCSD domains.

-!- SIMILARITY: Contains 1 SH3 domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U33058; AAB0542.1; -
EMBL; AF003131; AAB54132.2; -
PDB; 1FHO; 20-DEC-00.
WormPep; C09D1.1; CE30426.
InterPro; IPR008957; FN_III-like.
InterPro; IPR003961; FN_III.
InterPro; IPR007110; IG-like.
InterPro; IPR003598; IG_C2.
InterPro; IPR003006; IG_MHC.
InterPro; IPR001849; PH.
InterPro; IPR007850; RCSD.
InterPro; IPR000219; RhogEF.
InterPro; IPR001452; SH3.
Pfam; PF00041; fn3; 1.
Pfam; PF00047; ig; 47.
Pfam; PF00169; PH; 1.
Pfam; PF05177; RCSD; 5.
Pfam; PF00621; RhogEF; 1.
Pfam; PF00018; SH3; 1.
SMART; SM00408; IGc2; 23.
SMART; SM00325; RhogEF; 1.
SMART; SM00326; SH3; 1.
PROSITE; PS50010; DH_2; 1.
PROSITE; PS50835; IG_LIKE; 49.
PROSITE; PS50003; PH_DOMAIN; 1.
PROSITE; PS50002; SH3; 1.
Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;
3D-structure.

63 127 SH3.
152 330 DH.
342 498 FH.
547 633 IG-LIKE C2-TYPE 1.
648 736 IG-LIKE C2-TYPE 2.
748 838 IG-LIKE C2-TYPE 3.
946 1033 IG-LIKE C2-TYPE 4.
1044 1132 IG-LIKE C2-TYPE 5.
1140 1227 IG-LIKE C2-TYPE 6.
1272 1315 THR-RICH.
1375 1475 RCSD 1.
1479 1585 RCSD 2.
1597 1685 RCSD 3.
1700 1799 RCSD 4.
1800 1860 RCSD 5.
1982 2067 IG-LIKE C2-TYPE 7.
2071 2163 IG-LIKE C2-TYPE 8.
2171 2281 IG-LIKE C2-TYPE 9.
2269 2359 IG-LIKE C2-TYPE 10.
2455 2555 IG-LIKE C2-TYPE 11.
2463 2564 IG-LIKE C2-TYPE 12.
2563 2651 IG-LIKE C2-TYPE 13.
2657 2746 IG-LIKE C2-TYPE 14.
2754 2858 IG-LIKE C2-TYPE 15.
2887 2980 IG-LIKE C2-TYPE 16.

FT DOMAIN 2994 3081 IG-LIKE C2-TYPE 17.
FT DOMAIN 3087 3183 IG-LIKE C2-TYPE 18.
FT DOMAIN 3189 3280 IG-LIKE C2-TYPE 19.
FT DOMAIN 3286 3376 IG-LIKE C2-TYPE 20.
FT DOMAIN 3384 3469 IG-LIKE C2-TYPE 21.
FT DOMAIN 3482 3572 IG-LIKE C2-TYPE 22.
FT DOMAIN 3580 3667 IG-LIKE C2-TYPE 23.
FT DOMAIN 3686 3777 IG-LIKE C2-TYPE 24.
FT DOMAIN 3817 3908 IG-LIKE C2-TYPE 25.
FT DOMAIN 3920 4009 IG-LIKE C2-TYPE 26.
FT DOMAIN 4018 4106 IG-LIKE C2-TYPE 27.
FT DOMAIN 4109 4201 IG-LIKE C2-TYPE 28.
FT DOMAIN 4212 4297 IG-LIKE C2-TYPE 29.
FT DOMAIN 4302 4387 IG-LIKE C2-TYPE 30.
FT DOMAIN 4400 4485 IG-LIKE C2-TYPE 31.
FT DOMAIN 4489 4580 IG-LIKE C2-TYPE 32.
FT DOMAIN 4588 4678 IG-LIKE C2-TYPE 33.
FT DOMAIN 4681 4771 IG-LIKE C2-TYPE 34.
FT DOMAIN 4873 4961 IG-LIKE C2-TYPE 35.
FT DOMAIN 4965 5057 IG-LIKE C2-TYPE 36.
FT DOMAIN 5067 5160 IG-LIKE C2-TYPE 37.
FT DOMAIN 5171 5260 IG-LIKE C2-TYPE 38.
FT DOMAIN 5277 5366 IG-LIKE C2-TYPE 39.
FT DOMAIN 5383 5472 IG-LIKE C2-TYPE 40.
FT DOMAIN 5487 5578 IG-LIKE C2-TYPE 41.
FT DOMAIN 5595 5685 IG-LIKE C2-TYPE 42.
FT DOMAIN 5701 5790 IG-LIKE C2-TYPE 43.
FT DOMAIN 5815 5904 IG-LIKE C2-TYPE 44.
FT DOMAIN 5925 6014 IG-LIKE C2-TYPE 45.
FT DOMAIN 6038 6130 IG-LIKE C2-TYPE 46.
FT DOMAIN 6150 6239 IG-LIKE C2-TYPE 47.
FT DOMAIN 6275 6368 FIBRONECTIN TYPE-III.
FT DOMAIN 6413 6502 IG-LIKE C2-TYPE 48.
FT DOMAIN 6507 6596 IG-LIKE C2-TYPE 49.
FT DISULFID 568 621 POTENTIAL.
FT DISULFID 2908 2975 POTENTIAL.
FT DISULFID 3015 3085 POTENTIAL.
FT DISULFID 3707 3759 POTENTIAL.
FT DISULFID 3826 3890 POTENTIAL.
FT DISULFID 5092 5157 POTENTIAL.
FT DISULFID 5298 5350 POTENTIAL.
FT DISULFID 5508 5560 POTENTIAL.
FT DISULFID 5616 5669 POTENTIAL.
FT DISULFID 5722 5764 POTENTIAL.
FT DISULFID 5836 5901 POTENTIAL.
FT DISULFID 5946 5998 POTENTIAL.
FT DISULFID 6036 6171 POTENTIAL.
FT DISULFID 6421 6486 POTENTIAL.
FT DISULFID 2137 2137 A -> P (IN REF. 1).
FT CONFLICT 2245 2247 AKA -> PKP (IN REF. 1).
FT CONFLICT 2258 2258 A -> P (IN REF. 1).
FT CONFLICT 2284 2284 E -> G (IN REF. 1).
FT CONFLICT 2297 2297 M -> I (IN REF. 1).
FT CONFLICT 3531 3531 A -> G (IN REF. 1).
FT CONFLICT 3884 3888 DAGGY -> RRRRI (IN REF. 1).
FT CONFLICT 3929 3929 A -> V (IN REF. 1).
FT CONFLICT 5134 5134 A -> P (IN REF. 1).
FT CONFLICT 5145 5145 T -> S (IN REF. 1).
FT CONFLICT 5185 5185 G -> A (IN REF. 1).
FT CONFLICT 5199 5199 K -> N (IN REF. 1).
FT CONFLICT 5202 5202 L -> F (IN REF. 1).
FT CONFLICT 5213 5213 P -> L (IN REF. 1).
FT CONFLICT 6178 6178 A -> G (IN REF. 1).
FT CONFLICT 6268 6268 K -> E (IN REF. 1).
SQ SEQUENCE 6632 AA; 731665 MW; 262D3EDD62960B89 CRC64;

Query Match 45.6%; Score 47; DB 1; Length 6632;
Best Local Similarity 45.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
:::|::|::|::|
Db 1353 RRVSFARPEELPKVIDSDRK 1372

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

DR EMBL; S73885; AAB32235.1; --
DR EMBL; AC004653; AAC17116.1; --
DR EMBL; BC010576; AAH10576.1; --
DR EMBL; X57435; CAA40683.1; --
DR PIR; I56893; I56893.
DR HSP; P22415; IAN4.
DR TRANSFAC; T00036; --
DR Genew; HGNC:11745; TFAP4.
DR MIM; 600743; --
DR GO; GO:0003705; F:RNA polymerase II transcription factor acti. .; TAS.
DR GO; GO:0003713; P:transcription co-activator activity; TAS.
DR GO; GO:0007275; P:development; TAS.
DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; TAS.
DR InterPro; IPR001092; HLH_basic.
DR SMART; SM0010; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
KW Transcription regulation; DNA-binding; Activator; Nuclear protein.
FT DNA BIND 48 60 BASIC DOMAIN.
FT DOMAIN 61 100 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT DOMAIN 99 120 LEUCINE-ZIPPER 1.
FT DOMAIN 151 179 LEUCINE-ZIPPER 2.
FT DOMAIN 193 222 GLN-RICH.
FT DOMAIN 225 244 PRO-RICH.
FT CONFLICT 218 218 Q -> H (IN REF. 2). CRC64;
SQ SEQUENCE 338 AA; 38725 MW; 540C00856596B83 CRC64;

Query Match 44.7%; Score 46; DB 1; Length 338;
Best Local Similarity 38.9%; Pred. No. 10;
Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 ORLYAQEELSNEVLPPP 18
::: |||: |||:
Db 210 EKLEREQQLRTQLLPPP 227

RESULT 22
11SB CUCMA STANDARD; PRT; 480 AA.
ID 11SB CUCMA
AC P13744
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 11S globulin beta subunit precursor.
OS Cucurbita maxima (Pumpkin) (Winter squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3661;
RN [1] _SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kurokawa Anakuri Nankin;
RX MEDLINE=8166744; PubMed=2450746;
RA Hayashi M., Mori H., Nishimura M., Akazawa T., Hara-Nishimura I.;
RT "Nucleotide sequence of cloned cDNA coding for pumpkin 11-S globulin
RT beta subunit."
RL Eur. J. Biochem. 172:627-632(1988).
RN [2]
RP SEQUENCE OF 22-30 AND 297-302.
RA Ohmiya M., Hara I., Mastubara H.;
RT "Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the
RT acidic and basic peptide chains and identification of a pyroglutamy
RT peptide chain."
RL Plant Cell Physiol. 21:157-167(1980).
CC -!- FUNCTION: This is a seed storage protein.
CC -!- SUBUNIT: Hexamer; each subunit is composed of an acidic and a
CC basic chain derived from a single precursor and linked by a
CC disulfide bond.
CC -!- SIMILARITY: Belongs to the 11S seed storage protein (globulins)

RESULT 21
TAP4 HUMAN STANDARD; PRT; 338 AA.
ID TAP4 HUMAN
AC Q01664; O60409;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Transcription factor AP-4 (Activating enhancer-binding protein 4).
GN TFAP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1] _SEQUENCE FROM N.A.
RP MEDLINE=95018629; PubMed=7933101;
RA Ou S.H., Garcia-Martinez L.F., Paulsen E.J., Gaynor R.B.;
RT "Role of flanking E box motifs in human immunodeficiency virus type 1
RT TATA element function."
RL J. Virol. 68:7188-7199(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Ricke D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,
RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,
RA Goodwin L., Bryant J., Tesmer J., Meincke L., Longmire J., White S.,
RA Ueng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Misra M.,
RA Deaven L.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
TISSUE=Brain;
MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uesdin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
RN [4]
RP SEQUENCE OF 18-338 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91065520; PubMed=2123466;
RA Hu Y.-F., Luescher B., Admon A., Mermoud N., Tian R.;
RT "Transcription factor AP-4 contains multiple dimerization domains
RT that regulate dimer specificity."
RL Genes Dev. 4:1741-1752(1990).
CC -!- FUNCTION: Transcription factor that activates both viral and
CC cellular genes by binding to the symmetrical DNA sequence
CC 5'-CAGTGTG-3'.
CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another
CC bHLH protein, Homodimer.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M36407; AAA33110.1; --

InterPro; IPR006045; Cupin.

InterPro; IPR007113; Cupin sup.

InterPro; IPR006044; Seedstore_11s.

Pfam; PF00190; Cupin; 2.

PRINTS; PR00439; 11SGLOBULIN.

PROSITE; PS00305; 11S_SEED_STORAGE; 1.

Seed storage protein; Signal; Pyrrolidone carboxylic acid.

FT SIGNAL 1 21

FT CHAIN 22 480 11S GLOBULIN BETA SUBUNIT.

FT CHAIN 22 296 GAMMA CHAIN (ACIDIC).

FT CHAIN 297 480 DELTA CHAIN (BASIC).

FT MOD RES 22 22 PYRROLIDONE CARBOXYLIC ACID.

FT DISULFID 124 303 INTERCHAIN (GAMMA-DELTA) (POTENTIAL).

FT CONFLICT 27 27 S -> E (IN REF. 2).

FT CONFLICT 30 30 E -> S (IN REF. 2).

SEQUENCE 480 AA; 54625 MW; BCD8A83DD1AED93C CRC64;

Query Match 44.7%; Score 46; DB 1; Length 480;

Best Local Similarity 57.9%; Pred. No. 15;

Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPPP 19

Db 458 QRLKYQEQEM--RVLSPCR 474

RESULT 23

HEXA BLADI

ID HEXA BLADI STANDARD; PRT; 733 AA.

AC Q17127;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Hexamerin precursor.

OS Blaberus discoidalis (Tropical cockroach).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;

OC Blaberidae; Blaberus.

OC NCBI_TaxID=6981;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pat body;

RA Jamroz R.C., Beintema J.J., Stam W.T., Bradfield J.Y.;

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Larval storage protein (LSP) which may serve as a store of amino acids for synthesis of adult proteins (By similarity).

CC -!- SUBUNIT: Homohexamer (Potential).

CC -!- SUBCELLULAR LOCATION: Extracellular (By similarity).

CC -!- SIMILARITY: TO ARYLPHORINS AND TO ARTHROPOD HEMOCYANINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U01328; AAA74579.1; --

DR HSSP; P04253; LOXY.

DR InterPro; IPR008922; Di-copper centre.

DR InterPro; IPR008936; Hemocyanin.

DR InterPro; IPR005203; hemocyanin_C.

DR InterPro; IPR005204; hemocyanin_N.

DR InterPro; IPR007110; Ig-like.

DR Pfam; PF00372; hemocyanin; 1.

DR Pfam; PF03723; hemocyanin_C; 1.

DR Pfam; PF03722; hemocyanin_N; 1.

DR PRINTS; PR00187; HAEMOCYANIN.

DR PROSITE; PS00209; HEMOCYANIN 1; FALSE NEG.

DR PROSITE; PS00210; HEMOCYANIN 2; FALSE NEG.

KW Signal; Storage protein; Glycoprotein.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 733 HEXAMERIN.

FT CARBOHYD 199 199 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 431 431 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 733 AA; 87813 MW; 083DF739DD665729 CRC64;

Query Match 43.7%; Score 45; DB 1; Length 733;

Best Local Similarity 50.0%; Pred. No. 34;

Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 5 YAQEELSNEVLPPPP 18

Db 148 FQEDLSYILPPPP 161

RESULT 24

RPB2 ARATH

ID RPB2 ARATH STANDARD; PRT; 1188 AA.

AC P38420; Q9SVS6;

DT 01-OCT-1994 (Rel. 30, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE DNA-directed RNA polymerase II 135 kDa polypeptide (EC 2.7.7.6)

DE (RNA polymerase II subunit 2).

DE NCBI_TaxID=3702;

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OC NCBI_TaxID=3702;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=cv. Columbia;

RK MEDLINE=20083488; PubMed=10617198;

RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G., Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terry N., Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M., Mueller M., Weichselgartner M., de Simone V., Obermaier B., Maché R., Watson M., Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtheini T., Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I., Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P., Langham S.-A., McCullagh B., Bilham L., Robben J., Van der Schueren B., Grymonprez B., Chuang Y.-J., Vandenbussche F., Braeken M., Weijens I., Voet M., Bastiaens I., Aert R., Defoor E., Weitzenecker T., Bothe G., Rampsperger U., Hilbert H., Braun M., Holzer E., Brandt A., Peters S., van Staveren M., Dirks W., Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P., Bernsner S., Hempel S., Feldpausch M., Lambirth S., Van den Daele H., De Keyser A., Buyschaert C., Gielen J., Quail M.A., Bray-Allen S., Van Montagu M., Rogers J., Cronin A., Quail M.A., Mayes R., Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R., Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S., Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H., Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,

FT ZN_PING 1124 1145 C4-TYPE (POTENTIAL).
FT VARIANT 787 787 I -> N.
FT CONFLICT 354 355 PH -> LY (IN REF. 1).
SQ SEQUENCE 1188 AA; 135018 MW; C304E43515C2C364 CRC64;
Query Match 43.7%; Score 45; DB 1; Length 1188;
Best Local Similarity 50.0%; Pred. No. 59; Indels 0;
Matches 8; Conservative 5; Mismatches 3; Gaps 0;
QY 1 ORLKYAQEELSNEVLP 16
:|||||:|:|:
DB 339 KRIKYARDILQKEMLP 354
:|||||:|:|:
RESULT 25
PMG3 YEAST STANDARD; PRT; 303 AA.
ID PMG3 YEAST STANDARD; PRT; 303 AA.
AC Q12326;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoglycerate mutase 3 (EC 5.4.2.1) (phosphoglyceromutase 3) (PGAM
DE 3) (MPEG 3) (BPG-dependent PGAM 3).
GN GPM3 OR YOL056W OR Q1236.
OS Saccharomycetes cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / FY73;
RX MEDLINE=963181248; PubMed=8789261;
RA Manhaupt G., Vetter I., Schwarzlose C., Mittel S., Feldmann H.;
RT "Analysis of a 26 kb region on the left arm of yeast chromosome XV.";
RL Yeast 12:67-76(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Ansorge W., Benes V., Rechmann S., Schwager C., Teodoru C., Voss H.,
RA Wiemann S.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=98205882; PubMed=9544241;
RA Heinsch J.J., Mueller S., Schlueter E., Jacoby J., Rodicio R.;
RT "Investigation of two yeast genes encoding putative isoenzymes of
RT phosphoglycerate mutase.";
RL Yeast 14:203-213(1998).
CC -!- FUNCTION: Could be non-functional.
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate + 2,3-diphosphoglycerate
CC = 3-phospho-D-glycerate + 2,3-diphosphoglycerate.
CC -!- PATHWAY: Glycolysis.
CC -!- SIMILARITY: Belongs to the phosphoglycerate mutase family. BPG-
CC dependent PGAM subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X91067; CAA62530.1; -.
DR EMBL; Z74798; CAA99064.1; -.
DR PIR; S61723; S61723.
DR HSP; P00950; 5PGM.
DR GeneOnline; 143479; -.
DR SGD; S0005417; GPM3.
DR GO; GO:0004619; F:phosphoglycerate mutase activity; IMP.
DR InterPro; IPR001345; PG/BPGM mutase.
DR InterPro; IPR005952; Phosphogly_mut1.
DR Pfam; PF00300; PGAM; 1.
DR TIGRFAme; TIGR01258; pgm_1; 1.

RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
RA Neumann S., Argirion A., Vitale D., Liguori R., Piravandi E.,
RA Massenet O., Quigley F., Clabaud G., Muendlein A., Feiber R.,
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
RA Chedof F., Cooke R., Berger C., Monfort A., Casacuberta E.,
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
RA Perez-Perez A., Fumelle B., Bent E., Johnson S., Tacon D., Jesse T.,
RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
RA Frishman D., Haase D., Lencke K., Mewes H.-W., Stocker S.,
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
RA Sekhon M., Murray J., Sheel P., Cordes M., Abu-Threiden J.,
RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
RA Kramer J., Spiech J., Ryan E., Andrews S., Geisel C., Hillier L.W.,
RA Nelson J., Spiech J., Ryan E., Andrews S., Geisel C., Hillier L.W.,
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Jounu C.,
RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
RA Chen E., Maria M.A., Martensen R., McCombie W.R.;
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:769-777(1999).
CC -!- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription
CC of DNA into RNA using the four ribonucleoside triphosphates as
CC substrates.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -!- SUBUNIT: Each class of RNA polymerase is assembled from 9 to 14
CC different polypeptides. This subunit is the second largest
CC component of RNA polymerase II.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are
CC found in eukaryotic nuclei: polymerase I for the ribosomal RNA
CC precursor, polymerase II for the mRNA precursor, and polymerase
CC III for SS and tRNA genes.
CC -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z19120; CAA79527.1; -.
DR EMBL; Z19121; CAA79528.1; -.
DR EMBL; AL035527; CAB36815.1; -.
DR EMBL; AL161555; CAB81278.1; -.
DR PIR; T05846; T05846.
DR InterPro; IPR007121; RNA_pol_B.
DR InterPro; IPR007644; RNA_pol_Rpb2_1.
DR InterPro; IPR007642; RNA_pol_Rpb2_2.
DR InterPro; IPR007645; RNA_pol_Rpb2_3.
DR InterPro; IPR007646; RNA_pol_Rpb2_4.
DR InterPro; IPR007647; RNA_pol_Rpb2_5.
DR InterPro; IPR007120; RNA_pol_Rpb2_6.
DR InterPro; IPR007641; RNA_pol_Rpb2_7.
DR Pfam; PF04563; RNA_pol_Rpb2_1; 1.
DR Pfam; PF04561; RNA_pol_Rpb2_2; 1.
DR Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR Pfam; PF04566; RNA_pol_Rpb2_4; 1.
DR Pfam; PF04567; RNA_pol_Rpb2_5; 1.
DR Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Transferase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT DOMAIN 14 20 ASP/GLU-RICH (ACIDIC).

DR PROSITE, PS00175; PG MUTASE; 1.
KW Isomerase; Glycolysis.
FT ACT_SITE 14 14 TELP-PHOSPHOHLISTIDINE INTERMEDIATE
FT ACT_SITE 70 70 (BY SIMILARITY).
FT ACT_SITE 235 235 REQUIRED FOR BINDING CARBOXYL GROUP OF
FT ACT_SITE 303 AA; 34863 MW; 29C3FF3D28560914 CRC64;
SQ SEQUENCE 303 AA; 34863 MW; 29C3FF3D28560914 CRC64;
Query Match 42.7%; Score 44; DB 1; Length 303;
Best Local Similarity 56.2%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 1 QRLKYAQBELSNEVLP 16
DB 183 RHLKTYGPEEKANRLP 198
RESULT 26
HSLU_YERPE
ID HSLU_YERPE STANDARD; PRT; 443 AA.
AC Q8ZJ5;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP-dependent hsl protease ATP-binding subunit hslU.
GN HSLU OR YP00105 OR Y0294.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Pretice M.B., Sebatia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M.A., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner P.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
CC -!- FUNCTION: Chapterone subunit of a proteasome-like degradation
CC complex (By similarity).
CC -!- SUBUNIT: A double ring-shaped homohexamer of hslV is capped on
CC each side by a ring-shaped hslU homohexamer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the clpX chapterone family. HslU subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ414141; CAC89970.1; -.
CC EMBL; AB013628; AM83886.1; -.
CC PIR; AH0013; AH0013.

DR HAMAP; MF_00249; -; 1.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003959; AAA_ATPase_cent.
DR InterPro; IPR004491; Hsp_HslU.
DR Pfam; PF00004; AAA; 1.
DR SMART; SM00382; AAA; 1.
DR TIGRFAMs; TIGR00390; hslU; 1.
KW Chapterone; ATP-binding; Complete proteome.
FT NP_BIND 57 64 ATP (POTENTIAL).
SQ SEQUENCE 443 AA; 49804 MW; C8F480057C17E789 CRC64;
Query Match 42.7%; Score 44; DB 1; Length 443;
Best Local Similarity 41.7%; Pred. No. 28;
Matches 10; Conservative 5; Mismatches 5; Indels 4; Gaps 1;
QY 1 QRLKYAQBELSNE-----VLPFPRK 20
DB 117 EKMRYAEELABERILDVLPPAK 140
RESULT 27
GPMI ANTSP
ID GPMI ANTSP STANDARD; PRT; 510 AA.
AC Q06454;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
DE (EC 5.4.2.1) (Phosphoglyceromutase) (BPG-independent PGAM) (IPGM).
GN GPMI OR PGMA.
OS Antithamnion sp.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Florideophyceae; Ceramiales; Ceramiaceae;
OC Antithamnion.
OX NCBI_TaxID=2767;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93173098; PubMed=8437571;
RA Valentin K.-U.;
RT "SecA is plastid-encoded in a red alga: implications for the evolution
RT of plastid genomes and the thylakoid protein import apparatus.";
RL Mol. Gen. Genet. 236:245-250(1993).
CC -!- FUNCTION: Catalyzes the interconversion of 2-phosphoglycerate and
CC 3-phosphoglycerate (By similarity).
CC -!- CATALYTIC ACTIVITY: 2-phospho-D-glycerate = 3-phospho-D-glycerate.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- PATHWAY: Glycolysis.
CC -!- SUBCELLULAR LOCATION: Chloroplast.
CC -!- SIMILARITY: Belongs to the BPG-independent phosphoglycerate mutase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X64705; CAA45959.1; -.
CC PIR; S42705; S42705.
DR HAMAP; MF_01038; -; 1.
DR InterPro; IPR006124; Metalloenzyme.
DR InterPro; IPR005995; Pgm_bpd_ind.
DR Pfam; PF01676; Metalloenzyme_1.
DR ProDom; PD004429; Pgm_bpd_ind; 1.
DR TIGRFAMs; TIGR01307; Pgm_bpd_ind; 1.
KW Isomerase; Glycolysis; Metal-binding; Manganese; Chloroplast.
FT ACT_SITE 65 65 PHOSPHOSERINE INTERMEDIATE
FT ACT_SITE 65 65 (BY SIMILARITY).
FT METAL 15 15 MANGANESE 2 (BY SIMILARITY).
FT METAL 65 65 MANGANESE 2 (BY SIMILARITY).
FT METAL 399 399 MANGANESE 1 (BY SIMILARITY).

FT METAL 403 403 MANGANESE 1 (BY SIMILARITY).
FT METAL 440 440 MANGANESE 2 (BY SIMILARITY).
FT METAL 441 441 MANGANESE 2 (BY SIMILARITY).
FT METAL 458 458 MANGANESE 1 (BY SIMILARITY).
SQ SEQUENCE 510 AA; 57840 MW; 03DC7FD93A196B33 CRC64;

Query Match 42.7%; Score 44; DB 1; Length 510;
Best Local Similarity 43.8%; Pred. No. 33;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 4 KYAQEELSNEVLPPPRK 19
DB 226 KYEQNISDEFLPPTK 241
:::|||||:
:::|||||:

RESULT 28
ORCL1 YEAST STANDARD; PRT; 914 AA.
AC P54784;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Origin recognition complex subunit 1 (Origin recognition complex
protein 120 kDa subunit).
DE ORCL1 OR YML055W
GN Saccharomyces cerevisiae (Baker's yeast).
OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 381-408; 532-550; 642-653 AND
726-732.
RX MEDLINE=96069857; PubMed=7585959;
RA Bell S.P., Mitchell J., Leber J., Kobayashi R., Stillman B.;
RT "The multidomain structure of Orc1p reveals similarity to regulators
of DNA replication and transcriptional silencing.";
RL Cell 83:563-568 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96059484; PubMed=7579692;
RA Loo S., Fox C.A., Rine J., Kobayashi R., Stillman B., Bell S.P.;
RT "The origin recognition complex in silencing, cell cycle progression,
and DNA replication.";
RL Mol. Biol. Cell 6:741-756 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C / AB972;
RX MEDLINE=97313266; PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
Rice P., Skelton J., Walsh S., Whitehead S., Barrett B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII.";
RL Nature 387:90-93 (1997).
CC -!- FUNCTION: COMPONENT OF THE ORIGIN RECOGNITION COMPLEX (ORC) THAT
CC BINDS ORIGINS OF REPLICATION. IT HAS A ROLE IN BOTH CHROMOSOMAL
CC REPLICATION AND MATING TYPE TRANSCRIPTIONAL SILENCING. BINDS TO
CC THE ARS CONSENSUS SEQUENCE (ACS) OF ORIGINS OF REPLICATION.
CC -!- SUBUNIT: ORC IS COMPOSED OF SIX SUBUNITS OF 120 kDa, 71 kDa, 62
CC kDa, 56 kDa, 53 kDa AND 50 kDa.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: THE N-TERMINUS IS DEDICATED TO MATING-TYPE REPRESSION.
CC -!- SIMILARITY: Belongs to the ORC1 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U34860; AAB38248.1; -.
DR EMBL; Z38114; CA86256.1; -.
DR PIR; S48333; S48333.
DR PDB; 1W4Z; 11-SEP-02.
DR GerMOnline; 142596; -.
DR GK; P54784; -.
DR SGD; S0004530; ORC1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003959; AAA_ATPase_centr.
DR InterPro; IPR001025; BAH.
DR Pfam; PF00004; AAA; 1.
DR Pfam; PF01426; BAH; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00439; BAH; 1.
KW DNA replication; Nuclear protein; DNA-binding; ATP-binding;
KW 3D-structure.
FT NP_BIND 479 486 ATP (POTENTIAL).
FT NP_BIND 726 733 ATP (POTENTIAL).
FT DOMAIN 281 292 POLY-GLU.
FT DOMAIN 759 768 POLY-ASP.
SQ SEQUENCE 914 AA; 104399 MW; BDAA9D3E40B4B76A CRC64;

Query Match 42.7%; Score 44; DB 1; Length 914;
Best Local Similarity 30.0%; Pred. No. 63;
Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 310 RKIKLGKDDIDASVQPPPRK 329
:::|||||:
:::|||||:

RESULT 29
DMD_MOUSE STANDARD; PRT; 3678 AA.
ID DMD_MOUSE STANDARD;
AC P11531; O35653; Q60703;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Dystrophin.
GN DMD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92253376; PubMed=1579466;
RA Bies R.D., Phelps S.F., Cortez M.D., Roberts R., Caskey C.T.,
RA Chamberlain J.S.;
RT "Human and murine dystrophin mRNA transcripts are differentially
RT expressed during skeletal muscle, heart, and brain development.";
RL Nucleic Acids Res. 20:1725-1731 (1992).
RN [2]
RP SEQUENCE OF 1-201 FROM N.A.
RX MEDLINE=87273512; PubMed=3607877;
RA Koenig M., Hoffman E.P., Bertelson C.J., Monaco A.P., Feener C.,
RA Kunkel L.M.;
RT "Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and
RT preliminary genomic organization of the DMD gene in normal and
RT affected individuals.";
RL Cell 50:509-517 (1987).
RN [3]
RP SEQUENCE OF 120-176 FROM N.A.
RX STRAIN=129/J;
RX MEDLINE=92182520; PubMed=1543903;
RA Maconochie M.K., Brown S.D.M., Greenfield A.J.;
RT "Sequence analysis of two exons from the murine dystrophin locus.";
RL Mamm. Genome 2:64-68 (1992).
RN [4]
RP SEQUENCE OF 300-1390 FROM N.A.
RX MEDLINE=88018015; PubMed=3659917;
RA Hoffman E.P., Monaco A.P., Feener C.C., Kunkel L.M.;
RT "Conservation of the Duchenne muscular dystrophy gene in mice and

humans.";
Science 238:347-350(1987).
[5]
SEQUENCE OF 986-1056 FROM N.A.
STRAIN=CS7BL/10; TISSUE=Skeletal muscle;
MEDLINE=94154933; PubMed=8111539;
Chamberlain J.S., Phelps S.F., Cox G.A., Maichele A.J.,
Greenwood A.D.;
"PCR analysis of muscular dystrophy in mdx mice.";
Mol. Cell Biol. Hum. Dis. Ser. 3:167-189(1993).
[6]
ALTERNATIVE SPLICING.
STRAIN=CS7BL/10; TISSUE=Retina;
MEDLINE=95360002; PubMed=7533443;
DA Souza V.N., Nguyen T.M., Morris G.E., Karges W., Pillers D.-A.M.,
Ray P.N.;
"A novel dystrophin isoform is required for normal retinal
electrophysiology.";
Hum. Mol. Genet. 4:837-842(1995).
[7]
INTERACTION WITH SNTA1.
MEDLINE=9603613; PubMed=7547961;
Madhavan R., Jarrett H.W.;
"Interactions between dystrophin glycoprotein complex proteins.";
Biochemistry 34:12204-12209(1995).
[8]
INTERACTION WITH SNTB1.
MEDLINE=97362062; PubMed=9214383;
Peters M.F., Adams M.E., Froehner S.C.;
"Differential association of syntrophin pairs with the dystrophin
complex.";
J. Cell Biol. 138:81-93(1997).
CC - FUNCTION: May play a role in anchoring the cytoskeleton to the
plasma membrane.
CC - SUBUNIT: Interacts with the syntrophins SNTA1, SNTB1, SNTB2, SNTG1
and SNTG2.
CC - ALTERNATIVE PRODUCTS:
Event-Alternative splicing; Named isoforms=1;
Comment=At least 11 isoforms are produced;
Name=1;
IsoId=P11531-1; Sequence=Displayed;
CC - TISSUE SPECIFICITY: Differentially expressed during skeletal
muscle, heart, and brain development. Also expressed in retina.
CC - SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN, ABP-
120, ABP-180, OR BETA-FODRIN).
CC - SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC - SIMILARITY: Contains 22 spectrin repeats.
CC - SIMILARITY: Contains 1 WW domain.
CC - SIMILARITY: Contains 1 ZZ-type zinc finger.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; M68859; AAB02797.1; -;
DR EMBL; X58153; CAA41157.1; -;
DR EMBL; M18025; AAA37530.1; -;
DR EMBL; U56724; AAB01216.1; -;
DR EMBL; U15218; AAA87068.1; -;
DR PIR; S28916; S28916.
DR HSSP; P46939; IQAG.
DR MGD; MGI:94909; Dmd.
DR GO; GO:0045202; C:synaptic junction; IDA.
DR InterPro; IPR001589; Actbind actnin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001202; WW_Rsp5_WWP.
DR InterPro; IPR000433; Znf_ZZ.

DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; spectrin; 20.
DR Pfam; PF00397; WW; 1.
DR Pfam; PF00569; ZZ; 1.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 21.
DR SMART; SM00456; WW; 1.
DR SMART; SM00291; Znf_ZZ; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS00021; CH; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS00020; WW_DOMAIN_2; 1.
DR PROSITE; PS01357; ZF_ZZ_1; 1.
DR PROSITE; PS01357; ZF_ZZ_2; 1.
DR Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
KW Repeat; Alternative splicing; Zinc-finger.
FT DOMAIN 1 240
FT DOMAIN 15 119 CH 1.
FT DOMAIN 134 237 CH 2.
FT REPEAT 341 449 SPECTRIN 1.
FT REPEAT 450 558 SPECTRIN 2.
FT REPEAT 561 669 SPECTRIN 3.
FT REPEAT 721 830 SPECTRIN 4.
FT REPEAT 832 936 SPECTRIN 5.
FT REPEAT 945 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.
FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2106 2210 SPECTRIN 15.
FT REPEAT 2213 2318 SPECTRIN 16.
FT REPEAT 2468 2570 SPECTRIN 17.
FT REPEAT 2573 2679 SPECTRIN 18.
FT REPEAT 2682 2795 SPECTRIN 19.
FT REPEAT 2798 2900 SPECTRIN 20.
FT REPEAT 2902 2924 SPECTRIN 21.
FT REPEAT 2927 3033 SPECTRIN 22.
FT DOMAIN 3048 3081 WW.
FT ZN_FING 3300 3347 ZZ-TYPE.
FT DOMAIN 3459 3511 BINDS_SNTB1 (BY SIMILARITY).
FT CONFLICT 463 463 D -> H (IN REF. 4).
FT CONFLICT 677 677 S -> F (IN REF. 4).
FT CONFLICT 2337 2337 V -> L (IN REF. 1; AAB02797).
SQ SEQUENCE 3678 AA; 425810 MW; 1D2E74CF7DB035EE CRC64;
Query Match 42.2%; Score 43.5; DB 1; Length 3678;
Best Local Similarity 55.8%; Pred. No. 3.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
Qy 3 LKYAQEELSNEVLPPPK 20
Db 698 VKHAQSELPP---PPPK 712

RESULT 30
DMD_CANFA
ID DMD_CANFA STANDARD; PRT; 3680 AA.
AC O37592;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dystrophin.
GN DMD.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OC NCBI_TaxID=9615;
RX [1]

RP SEQUENCE FROM N.A.
RC STRAIN=Golden retriever;
RA Carville K.S., Mann C.J., Schatzberg S.J., Wilton S.D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May play a role in anchoring the cytoskeleton to the
CC plasma membrane.
CC -!- SUBUNIT: Interacts with the syntrophins SNTA1, SNTB1, SNTB2, SNTG1
CC and SNTG2 (By similarity).
CC -!- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FIMBRIN,
CC ABP-120, ABP-180, OR BETA-FODRIN).
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 22 spectrin repeats.
CC -!- SIMILARITY: Contains 1 WW domain.
CC -!- SIMILARITY: Contains 1 ZZ-type zinc finger.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF070485; AAC83646.1; -.
DR HSPSP; P46939; IQAG.
DR InterPro; IPR001589; Actbind_actnin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001202; WW Rep5_WWP.
DR InterPro; IPR000433; Znf_ZZ.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00435; spectrin; 21.
DR Pfam; PF00397; WW; 1.
DR Pfam; PF00569; ZZ; 1.
DR SMART; SM00033; CH; 2.
DR SMART; SM00150; SPEC; 21.
DR SMART; SM00456; WW; 1.
DR SMART; SM00291; Znf_ZZ; 1.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; 1.
DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS01159; WW_DOMAIN_1; 1.
DR PROSITE; PS50020; WW_DOMAIN_2; 1.
DR PROSITE; PS01357; ZF_ZZ_1; 1.
DR PROSITE; PS50135; ZF_ZZ_2; 1.
KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
KW Repeat; Zinc-finger.
FT DOMAIN 15 119 CH 1.
FT DOMAIN 134 237 CH 2.
FT REPEAT 340 448 SPECTRIN 1.
FT REPEAT 449 557 SPECTRIN 2.
FT REPEAT 560 668 SPECTRIN 3.
FT REPEAT 720 829 SPECTRIN 4.
FT REPEAT 831 935 SPECTRIN 5.
FT REPEAT 944 1047 SPECTRIN 6.
FT REPEAT 1050 1156 SPECTRIN 7.
FT REPEAT 1159 1265 SPECTRIN 8.
FT REPEAT 1268 1369 SPECTRIN 9.
FT REPEAT 1470 1570 SPECTRIN 10.
FT REPEAT 1573 1678 SPECTRIN 11.
FT REPEAT 1681 1782 SPECTRIN 12.
FT REPEAT 1879 1981 SPECTRIN 13.
FT REPEAT 2013 2103 SPECTRIN 14.
FT REPEAT 2106 2210 SPECTRIN 15.
FT REPEAT 2213 2320 SPECTRIN 16.
FT REPEAT 2470 2572 SPECTRIN 17.
FT REPEAT 2575 2681 SPECTRIN 18.
FT REPEAT 2684 2797 SPECTRIN 19.
FT REPEAT 2800 2902 SPECTRIN 20.
FT REPEAT 2904 2926 SPECTRIN 21.
FT REPEAT 2929 3035 SPECTRIN 22.
FT DOMAIN 3050 3083 WW.

FT ZN_PING 3302 3349 ZZ-TYPE.
SQ SEQUENCE 3680 AA; 425650 MW; 539F1C9D72377872 CRC64;
Query Match 42.2%; Score 43.5; DB 1; Length 3680;
Best Local Similarity 55.6%; Pred. No. 3.5e+02;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
QY 3 LKQAQEELSNEVLPPPRK 20
Db 697 VKHAQEELPP---PPPOK 711
Search completed: May 11, 2004, 13:38:02
Job time : 8.75 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:28:26 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-09-171-432A-41

Perfect score: 103

Sequence: 1 QRLKYAQEELSNEVLPPPRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	82	12 Q8BDN1	Q8BDN1 hepatitis a
2	103	100.0	82	12 Q8BDN0	Q8BDN0 hepatitis a
3	103	100.0	82	12 Q8BDN9	Q8BDN9 hepatitis a
4	103	100.0	82	12 Q8BDM8	Q8BDM8 hepatitis a
5	103	100.0	82	12 Q8BDM7	Q8BDM7 hepatitis a
6	103	100.0	82	12 Q8BDM6	Q8BDM6 hepatitis a
7	103	100.0	82	12 Q8BDM5	Q8BDM5 hepatitis a
8	103	100.0	82	12 Q8BDM4	Q8BDM4 hepatitis a
9	103	100.0	82	12 Q8BDM3	Q8BDM3 hepatitis a
10	103	100.0	82	12 Q8BDM2	Q8BDM2 hepatitis a
11	103	100.0	82	12 Q8BDM1	Q8BDM1 hepatitis a
12	103	100.0	82	12 Q8BDM0	Q8BDM0 hepatitis a
13	103	100.0	82	12 Q8BDL9	Q8BDL9 hepatitis a
14	103	100.0	82	12 Q8BDL8	Q8BDL8 hepatitis a
15	103	100.0	82	12 Q8BDL7	Q8BDL7 hepatitis a
16	103	100.0	82	12 Q8BDL6	Q8BDL6 hepatitis a

17	103	100.0	82	12 Q8BDL5	Q8BDL5 hepatitis a
18	103	100.0	94	12 Q8ENT9	Q8ENT9 hepatitis a
19	103	100.0	94	12 Q8ENU3	Q8ENU3 hepatitis a
20	103	100.0	94	12 Q8ENV9	Q8ENV9 hepatitis a
21	103	100.0	94	12 Q8ENV5	Q8ENV5 hepatitis a
22	103	100.0	94	12 Q8ENV0	Q8ENV0 hepatitis a
23	103	100.0	94	12 Q8ENS8	Q8ENS8 hepatitis a
24	103	100.0	94	12 Q8ENT8	Q8ENT8 hepatitis a
25	103	100.0	94	12 Q8ENU2	Q8ENU2 hepatitis a
26	103	100.0	94	12 Q8ENT3	Q8ENT3 hepatitis a
27	103	100.0	94	12 Q8ENT1	Q8ENT1 hepatitis a
28	103	100.0	94	12 Q8ENT0	Q8ENT0 hepatitis a
29	103	100.0	94	12 Q8ENU1	Q8ENU1 hepatitis a
30	103	100.0	94	12 Q8ENT4	Q8ENT4 hepatitis a
31	103	100.0	94	12 Q8ENT2	Q8ENT2 hepatitis a
32	103	100.0	94	12 Q8ENU8	Q8ENU8 hepatitis a
33	103	100.0	94	12 Q8ENV0	Q8ENV0 hepatitis a
34	103	100.0	94	12 Q8ENT7	Q8ENT7 hepatitis a
35	103	100.0	94	12 Q8ENV2	Q8ENV2 hepatitis a
36	103	100.0	94	12 Q8ENU5	Q8ENU5 hepatitis a
37	103	100.0	94	12 Q8ENV1	Q8ENV1 hepatitis a
38	103	100.0	94	12 Q8ENV9	Q8ENV9 hepatitis a
39	103	100.0	94	12 Q8ENV7	Q8ENV7 hepatitis a
40	103	100.0	94	12 Q8ENV8	Q8ENV8 hepatitis a
41	103	100.0	94	12 Q8ENV6	Q8ENV6 hepatitis a
42	103	100.0	94	12 Q8ENW0	Q8ENW0 hepatitis a
43	103	100.0	94	12 Q8ENV6	Q8ENV6 hepatitis a
44	103	100.0	94	12 Q8ENU4	Q8ENU4 hepatitis a
45	103	100.0	94	12 Q8ENT6	Q8ENT6 hepatitis a
46	103	100.0	94	12 Q8ENW1	Q8ENW1 hepatitis a
47	103	100.0	94	12 Q8ENT5	Q8ENT5 hepatitis a
48	103	100.0	94	12 Q8ENV7	Q8ENV7 hepatitis a
49	103	100.0	94	12 Q8ENV4	Q8ENV4 hepatitis a
50	103	100.0	97	12 Q8QCR7	Q8QCR7 hepatitis a
51	103	100.0	97	12 Q8QCT0	Q8QCT0 hepatitis a
52	103	100.0	97	12 Q8QCS2	Q8QCS2 hepatitis a
53	103	100.0	97	12 Q8QCS6	Q8QCS6 hepatitis a
54	103	100.0	97	12 Q8QCR0	Q8QCR0 hepatitis a
55	103	100.0	97	12 Q8QCR6	Q8QCR6 hepatitis a
56	103	100.0	97	12 Q8QCS5	Q8QCS5 hepatitis a
57	103	100.0	97	12 Q8QCS0	Q8QCS0 hepatitis a
58	103	100.0	97	12 Q8QCS8	Q8QCS8 hepatitis a
59	103	100.0	97	12 Q8QCS3	Q8QCS3 hepatitis a
60	103	100.0	97	12 Q8QCS9	Q8QCS9 hepatitis a
61	103	100.0	97	12 Q8QCT1	Q8QCT1 hepatitis a
62	103	100.0	97	12 Q8QCS1	Q8QCS1 hepatitis a
63	103	100.0	97	12 Q8QCR3	Q8QCR3 hepatitis a
64	103	100.0	97	12 Q8QCS7	Q8QCS7 hepatitis a
65	103	100.0	97	12 Q8QCR1	Q8QCR1 hepatitis a
66	103	100.0	97	12 Q8QCC7	Q8QCC7 hepatitis a
67	103	100.0	97	12 Q8QCS6	Q8QCS6 hepatitis a
68	103	100.0	97	12 Q8QCS8	Q8QCS8 hepatitis a
69	103	100.0	97	12 Q8QCR2	Q8QCR2 hepatitis a
70	103	100.0	97	12 Q8QCR4	Q8QCR4 hepatitis a
71	103	100.0	97	12 Q8QCS5	Q8QCS5 hepatitis a
72	103	100.0	97	12 Q8QCS4	Q8QCS4 hepatitis a
73	103	100.0	97	12 Q8QCS9	Q8QCS9 hepatitis a
74	103	100.0	97	12 Q8QCR9	Q8QCR9 hepatitis a
75	103	100.0	97	12 Q8QCR8	Q8QCR8 hepatitis a
76	103	100.0	97	12 Q8QCR5	Q8QCR5 hepatitis a
77	103	100.0	97	12 Q8QCT2	Q8QCT2 hepatitis a
78	103	100.0	115	12 Q8DMR2	Q8DMR2 hepatitis a
79	103	100.0	115	12 Q8DMR6	Q8DMR6 hepatitis a
80	103	100.0	115	12 Q8DMR4	Q8DMR4 hepatitis a
81	103	100.0	115	12 Q8DMQ8	Q8DMQ8 hepatitis a
82	103	100.0	115	12 Q8DMR0	Q8DMR0 hepatitis a
83	103	100.0	115	12 Q8DMR5	Q8DMR5 hepatitis a
84	103	100.0	115	12 Q8DMQ9	Q8DMQ9 hepatitis a
85	103	100.0	115	12 Q8DMQ7	Q8DMQ7 hepatitis a
86	103	100.0	115	12 Q8DMR3	Q8DMR3 hepatitis a
87	103	100.0	116	12 Q82941	Q82941 hepatitis a
88	103	100.0	116	12 Q8W7X8	Q8W7X8 hepatitis a
89	103	100.0	116	12 Q8W7T4	Q8W7T4 hepatitis a

90 103 100.0 116 12 Q9W7X7 hepatitis a
91 103 100.0 116 12 Q9W7S7 hepatitis a
92 103 100.0 116 12 Q71977 hepatitis a
93 103 100.0 116 12 Q9W7T0 hepatitis a
94 103 100.0 116 12 Q71975 hepatitis a
95 103 100.0 116 12 Q71978 hepatitis a
96 103 100.0 116 12 Q8B8K6 hepatitis a
97 103 100.0 116 12 Q8B8K5 hepatitis a
98 103 100.0 116 12 Q8B8K4 hepatitis a
99 103 100.0 116 12 Q8B8K3 hepatitis a
100 103 100.0 116 12 Q8B8K2 hepatitis a

ALIGNMENTS

RESULT 1
Q8BDN1 PRELIMINARY; PRT; 82 AA.
AC Q8BDN1
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg283;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453451; AAN52259.1; -.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9505 MW; 8085191744E9C6E8 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 2
Q8BDN0 PRELIMINARY; PRT; 82 AA.
AC Q8BDN0
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg058;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453452; AAN52260.1; -.
FT NON_TER 1
FT NON_TER 82

SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 3
Q8BDM9 PRELIMINARY; PRT; 82 AA.
AC Q8BDM9
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg987;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453453; AAN52261.1; -.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 4
Q8BDM8 PRELIMINARY; PRT; 82 AA.
AC Q8BDM8
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg441;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453454; AAN52262.1; -.
FT NON_TER 1
FT NON_TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 5
Q8BDM7 PRELIMINARY; PRT; 82 AA.

AC Q8BDM7; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg799;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453455; AAN52263.1; --
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 6
Q8BDM6 PRELIMINARY; PRT; 82 AA.

AC Q8BDM6; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg002;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453456; AAN52264.1; --
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 7
Q8BDM5 PRELIMINARY; PRT; 82 AA.

AC Q8BDM5; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg139;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453457; AAN52265.1; --
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 8
Q8BDM4 PRELIMINARY; PRT; 82 AA.

AC Q8BDM4; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg927;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453458; AAN52266.1; --
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 9
Q8BDM3 PRELIMINARY; PRT; 82 AA.

AC Q8BDM3; (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
ON NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg927;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina."
RL J. Med. Virol. 68:168-174(2002).
DR EMBL; AF453459; AAN52267.1; --
FT NON_TER 1 1
FT NON_TER 82 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Arg548;
 RX MEDLINE=22198947; PubMed=12210404;
 RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
 RT "Genetic characterization of hepatitis A virus isolates from Buenos
 Aires, Argentina.";
 RL J. Med. Virol. 68:168-174(2002).
 DR EMBL: AF453459; AAN52267.1; -.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
 Query Match 100.0%; Score 103; DB 12; Length 82;
 Best Local Similarity 100.0%; Pred. No. 6.7e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 Db 61 QRLKYAQEELSNEVLPPPRK 80
 RESULT 10
 Q8BDM2 PRELIMINARY; PRT; 82 AA.
 AC Q8BDM2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Arg156;
 RX MEDLINE=22198947; PubMed=12210404;
 RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
 RT "Genetic characterization of hepatitis A virus isolates from Buenos
 Aires, Argentina.";
 RL J. Med. Virol. 68:168-174(2002).
 DR EMBL: AF453460; AAN52268.1; -.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
 Query Match 100.0%; Score 103; DB 12; Length 82;
 Best Local Similarity 100.0%; Pred. No. 6.7e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 Db 61 QRLKYAQEELSNEVLPPPRK 80
 RESULT 11
 Q8BDM1 PRELIMINARY; PRT; 82 AA.
 AC Q8BDM1;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Arg642;
 RX MEDLINE=22198947; PubMed=12210404;
 RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
 RT "Genetic characterization of hepatitis A virus isolates from Buenos
 Aires, Argentina.";
 RL J. Med. Virol. 68:168-174(2002).
 DR EMBL: AF453461; AAN52269.1; -.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;
 Query Match 100.0%; Score 103; DB 12; Length 82;
 Best Local Similarity 100.0%; Pred. No. 6.7e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 Db 61 QRLKYAQEELSNEVLPPPRK 80
 RESULT 12
 Q8BDM0 PRELIMINARY; PRT; 82 AA.
 AC Q8BDM0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Arg784;
 RX MEDLINE=22198947; PubMed=12210404;
 RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
 RT "Genetic characterization of hepatitis A virus isolates from Buenos
 Aires, Argentina.";
 RL J. Med. Virol. 68:168-174(2002).
 DR EMBL: AF453462; AAN52270.1; -.
 FT NON_TER 1
 FT NON_TER 82
 SQ SEQUENCE 82 AA; 9503 MW; 8083C4D698E9C6E8 CRC64;
 Query Match 100.0%; Score 103; DB 12; Length 82;
 Best Local Similarity 100.0%; Pred. No. 6.7e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QRLKYAQEELSNEVLPPPRK 20
 Db 61 QRLKYAQEELSNEVLPPPRK 80
 RESULT 13
 Q8BDL9 PRELIMINARY; PRT; 82 AA.
 AC Q8BDL9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Polyprotein (Fragment).
 OS Hepatitis A virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=Arg818;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina.";
RL J. Med. Virol. 68:168-174(2002).
DR EMBL: AF453463; AAN52271.1; --
FT NON TER 1
FT NON TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 14
Q8BDL8 PRELIMINARY; PRT; 82 AA.
AC Q8BDL8;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg873;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina.";
RL J. Med. Virol. 68:168-174(2002).
DR EMBL: AF453464; AAN52272.1; --
FT NON TER 1
FT NON TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 15
Q8BDL7 PRELIMINARY; PRT; 82 AA.
AC Q8BDL7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg143;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina.";

RL J. Med. Virol. 68:168-174(2002).
DR EMBL: AF453465; AAN52273.1; --
FT NON TER 1
FT NON TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 16
Q8BDL6 PRELIMINARY; PRT; 82 AA.
AC Q8BDL6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg081;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina.";
RL J. Med. Virol. 68:168-174(2002).
DR EMBL: AF453466; AAN52274.1; --
FT NON TER 1
FT NON TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
Db 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 17
Q8BDL5 PRELIMINARY; PRT; 82 AA.
AC Q8BDL5;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Arg578;
RX MEDLINE=22198947; PubMed=12210404;
RA Mbayed V.A., Sookoian S., Alfonso V., Campos R.H.;
RT "Genetic characterization of hepatitis A virus isolates from Buenos Aires, Argentina.";
RL J. Med. Virol. 68:168-174(2002).
DR EMBL: AF453467; AAN52275.1; --
FT NON TER 1
FT NON TER 82
SQ SEQUENCE 82 AA; 9521 MW; 808504D698E9C6E8 CRC64;

FT NON TER 94 94
SQ SEQUENCE 94 AA; 10892 MW; 9AF8BE91BB8C53CC CRC64;
Query Match 100.0%; Score 103; DB 12; Length 82;
Best Local Similarity 100.0%; Pred. No. 6.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
| | | | | | | | | | | | | | | | | | | | | |
DB 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 18
Q9ENT9 PRELIMINARY; PRT; 94 AA.
AC Q9ENT9
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 32;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038298; BAB11836.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON TER 94 94
SQ SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;
Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
| | | | | | | | | | | | | | | | | | | | | |
DB 61 QRLKYAQEELSNEVLPPPRK 80

RESULT 19
Q9ENU3 PRELIMINARY; PRT; 94 AA.
AC Q9ENU3
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 27;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038294; BAB11832.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON TER 94 94

RESULT 20
Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038282; BAB11820.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.

RESULT 21
Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038278; BAB11816.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.

DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4EAC CRC64;
Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 26
Q9ENT3 PRELIMINARY; PRT; 94 AA.
ID Q9ENT3
AC Q9ENT3; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 45;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038304; BAB11842.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10828 MW; 74184C635B86E1A6 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 27
Q9ENT1 PRELIMINARY; PRT; 94 AA.
ID Q9ENT1
AC Q9ENT1; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 49;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)

RT infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038306; BAB11844.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;
Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 28
Q9ENT0 PRELIMINARY; PRT; 94 AA.
ID Q9ENT0
AC Q9ENT0; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 53;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038307; BAB11845.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
SQ SEQUENCE 94 AA; 10858 MW; 9AF8BE91BB895FAC CRC64;
Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 61 QRLKYAQEELSNEVLPPPRK 80
RESULT 29
Q9ENU1 PRELIMINARY; PRT; 94 AA.
ID Q9ENU1
AC Q9ENU1; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Polypeptide (Fragment)
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 29;

```

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038296; BAB11834.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ORLYAQEELSNEVLPPRK 20
Db 61 ORLYAQEELSNEVLPPRK 80

RESULT 30
Q9ENT4 PRELIMINARY; PRT; 94 AA.
AC Q9ENT4;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
CX NCBI_TaxID=12092;
RN [1]_TaxID=12092;
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 44;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038303; BAB11841.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10828 MW; 74184C635BB6E1A6 CRC64;

Query Match 100.0%; Score 103; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 7.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 ORLYAQEELSNEVLPPRK 20
Db 61 ORLYAQEELSNEVLPPRK 80

Search completed: May 11, 2004, 13:40:24
Job time : 33 secs

```


OM protein - protein search, using sw model

Run on: May 11, 2004, 13:34:17 ; Search time 13.5 Seconds
(without alignments)
76.483 Million cell updates/sec

Title: US-09-171-432A-41

Perfect score: 103

Sequence: 1 QRLKYAQEELSNEVLPPPRK 20

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A COMB.pdp.*

2: /cgn2_6/ptodata/2/iaa/5B COMB.pdp.*

3: /cgn2_6/ptodata/2/iaa/6A COMB.pdp.*

4: /cgn2_6/ptodata/2/iaa/6B COMB.pdp.*

5: /cgn2_6/ptodata/2/iaa/PCFUS COMB.pdp.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	1091	6	Patent No. 5516630-2
2	103	100.0	2227	3	US-08-475-886-2
3	103	100.0	2227	3	US-08-475-886-5
4	103	100.0	2227	3	US-08-397-232-2
5	103	100.0	2227	3	US-08-397-232-4
6	103	100.0	2227	3	US-09-171-387-2
7	103	100.0	2227	4	US-09-653-499-2
8	103	100.0	2227	4	US-09-653-499-6
9	103	100.0	2227	4	US-10-104-966-12
10	103	100.0	2227	4	US-10-135-988-2
11	103	100.0	2227	4	US-10-135-988-6
12	98	95.1	2227	3	US-08-475-886-4
13	98	95.1	2227	4	US-09-653-499-4
14	98	95.1	2227	4	US-10-135-988-4
15	94	91.3	839	1	US-08-087-016-2
16	44	42.7	914	1	US-08-484-105-2
17	44	42.7	914	1	US-08-484-106-2
18	44	42.7	914	1	US-08-241-853-15
19	43.5	42.2	113	1	US-08-850-917-15
20	43.5	42.2	113	2	US-08-850-917-15
21	43	41.7	209	4	US-09-543-681A-7657
22	43	41.7	390	4	US-09-308-003-12
23	43	41.7	414	4	US-09-252-991A-22296
24	43	41.7	585	2	US-08-453-848-11
25	43	41.7	585	3	US-09-169-027-11
26	43	41.7	586	2	US-08-453-848-19
27	43	41.7	586	3	US-09-169-027-19

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BARCUDY, BAIGIE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091

5516630-2
Query Match 100.0%; Score 103; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORLKYAQEELSNEVLPPPRK 20
Db 1047 ORLKYAQEELSNEVLPPPRK 1066
|||||

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1993-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 103; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORLKYAQEELSNEVLPPPRK 20
Db 810 ORLKYAQEELSNEVLPPPRK 829
|||||

RESULT 3
US-08-475-886-6
; Sequence 6, Application US/08475886A

; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1993-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 103; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORLKYAQEELSNEVLPPPRK 20
Db 810 ORLKYAQEELSNEVLPPPRK 829
|||||

RESULT 4
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 103; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ORLKYAQEELSNEVLPPPRK 20
Db 810 ORLKYAQEELSNEVLPPPRK 829
|||||

RESULT 5
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 2227
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 103; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPRK 20
Db 810 QRLKYAQEELSNEVLPPRK 829

RESULT 6
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; FURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feller
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4229US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 103; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPRK 20
Db 810 QRLKYAQEELSNEVLPPRK 829

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: FURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 103; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPRK 20
Db 810 QRLKYAQEELSNEVLPPRK 829

RESULT 8
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: FURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 103; DB 4; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 9
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 103; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 10
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 103; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
|||||

Db 810 QRLKYAQEELSNEVLPPPRK 829
|||||

RESULT 11
US-10-135-988-6
; Sequence 6, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 103; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 12
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 95.1%; Score 98; DB 3; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 13

US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 95.1%; Score 98; DB 4; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 14
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 95.1%; Score 98; DB 4; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135

GENERAL INFORMATION:
; APPLICANT: NAINAN, OMANA V.
; APPLICANT: MARGOLIS, HAROLD S.
; APPLICANT: ROBERTSON, BETTY H.
; APPLICANT: BRINTON, MARGO H.
; APPLICANT: EBERT, JAMES W.
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1615 L Street N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,016
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,828
; FILING DATE: 03-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 839 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-087-016-2

Query Match 91.3%; Score 94; DB 1; Length 839;
Best Local Similarity 85.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 809 QRLKYAQEELSNEVLPPPRK 828

RESULT 16
US-09-252-991A-18339
; Sequence 18339, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18339
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18339

US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 95.1%; Score 98; DB 4; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 14
US-10-135-988-4
; Sequence 4, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 95.1%; Score 98; DB 4; Length 2227;
Best Local Similarity 95.0%; Pred. No. 8.5e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 15
US-08-087-016-2
; Sequence 2, Application US/08087016
; Patent No. 5430135

Query Match 42.7%; Score 44; DB 4; Length 180;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPK 19
Db 29 RLRYAAQRLQRLVLPQR 47
RESULT 17
US-08-484-105-2
Sequence 2, Application US/08484105
Patent No. 5589341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 914 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-105-2
Query Match 42.7%; Score 44; DB 1; Length 914;
Best Local Similarity 30.0%; Pred. No. 1e+02;
Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPK 20
Db 310 RKIKGKDDIDASVQPPPK 329
RESULT 19
US-08-241-853-15
Sequence 15, Application US/08241853
Patent No. 5693488
GENERAL INFORMATION:
APPLICANT: Fang, Kathy S.
APPLICANT: Hanafusa, Hidesaburo
TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:

GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 914 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-106-2
Query Match 42.7%; Score 44; DB 1; Length 914;
Best Local Similarity 30.0%; Pred. No. 1e+02;
Matches 6; Conservative 8; Mismatches 6; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPPPK 20
Db 310 RKIKGKDDIDASVQPPPK 329
RESULT 19
US-08-241-853-15
Sequence 15, Application US/08241853
Patent No. 5693488
GENERAL INFORMATION:
APPLICANT: Fang, Kathy S.
APPLICANT: Hanafusa, Hidesaburo
TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/241,853
FILING DATE: 12-MAY-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-078
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-241-853-15
Query Match 42.2%; Score 43.5; DB 1; Length 113;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
QY 3 LKVAQEELSNEVLPPPRK 20
DB 7 VKHAQEELPP---PPQK 21
RESULT 20
US-08-850-917-15
Sequence 15, Application US/08850917
Patent No. 5854045
GENERAL INFORMATION:
APPLICANT: Fang, Kathy S.
APPLICANT: Hanafusa, Hidesaburo
TITLE OF INVENTION: A TRANSMEMBRANE TYROSINE PHOSPHATASE
TITLE OF INVENTION: AND METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Klauber & Jackson
STREET: 411 Hackensack Avenue
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,917
FILING DATE: 02-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/241,853
FILING DATE: 12-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-078
TELEPHONE: 201 487-5800
TELEFAX: 201 343-1684
TELEX: 133521

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 113 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-850-917-15
Query Match 42.2%; Score 43.5; DB 2; Length 113;
Best Local Similarity 55.6%; Pred. No. 13;
Matches 10; Conservative 3; Mismatches 2; Indels 3; Gaps 1;
QY 3 LKVAQEELSNEVLPPPRK 20
DB 7 VKHAQEELPP---PPQK 21
RESULT 21
US-09-543-681A-7657
Sequence 7657, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7657
LENGTH: 209
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7657
Query Match 41.7%; Score 43; DB 4; Length 209;
Best Local Similarity 61.5%; Pred. No. 29;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 4 KYAQEELSNEVL 16
DB 100 KFALEKFSNELL 112
RESULT 22
US-09-308-003-12
Sequence 12, Application US/09308003
Patent No. 6326170
GENERAL INFORMATION:
APPLICANT: Burnham, Martin K. R.
APPLICANT: Lonetto, Michael A.
APPLICANT: Warren, Patrick V.
TITLE OF INVENTION: NOVEL PROKARYOTIC POLYNUCLEOTIDES,
TITLE OF INVENTION: POLYPEPTIDES AND THEIR USES
FILE REFERENCE: GM10093
CURRENT APPLICATION NUMBER: US/09/308,003
CURRENT FILING DATE: 1999-05-10
EARLIER APPLICATION NUMBER: 60/058,710
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 390
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-308-003-12
Query Match 41.7%; Score 43; DB 4; Length 390;

Best Local Similarity 57.1%; Pred. No. 58;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 3 LKVAQEELSNEVLPP 16
Db 368 IKVALKEASNESIP 381
RESULT 23
US-09-252-991A-22296
; Sequence 22296, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22296
; LENGTH: 414
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22296
Query Match 41.7%; Score 43; DB 4; Length 414;
Best Local Similarity 55.6%; Pred. No. 62;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVLPP 18
Db 342 QRLDYHAGELDLLELPP 359
RESULT 24
US-08-453-848-11
; Sequence 11, Application US/08453848
; Patent No. 5858368
; GENERAL INFORMATION:
; APPLICANT: Smith, Gale Eugene
; APPLICANT: Volvovitz, Franklin
; APPLICANT: Wilkinson, Bethanie Eident
; APPLICANT: Voznesensky, Andrei I.
; APPLICANT: Hackett, Craig Stanway
; TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
; TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,848
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/120,607
; FILING DATE: 13-SEPT-1993
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MGS101CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-873-8794
; TELEFAX: (404)-873-8795
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Influenza virus
; INDIVIDUAL ISOLATE: B/Panama/45/90 rHA
; FEATURE:
; NAME/KEY: HA signal peptide
; LOCATION: 1 to 17
; FEATURE:
; NAME/KEY: mature rHA
; LOCATION: 18 to 568
US-08-453-848-11
Query Match 41.7%; Score 43; DB 2; Length 585;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
QY 1 QRLKYAQEELSNEVL 15
Db 426 QRLSGAMDELHNEIL 440
RESULT 25
US-09-169-027-11
; Sequence 11, Application US/09169027
; Patent No. 6245532
; GENERAL INFORMATION:
; APPLICANT: Smith, Gale Eugene
; APPLICANT: Volvovitz, Franklin
; APPLICANT: Wilkinson, Bethanie Eident
; APPLICANT: Voznesensky, Andrei I.
; APPLICANT: Hackett, Craig Stanway
; TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
; TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/169,027
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/453,848
; FILING DATE: 30-MAY-1995
; APPLICATION NUMBER: 08/120,607
; FILING DATE: 13-SEPT-1993
; ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MGS101CIP
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 585 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Panama/45/90 rHA
FEATURE:
NAME/KEY: HA signal peptide
LOCATION: 1 to 17
FEATURE:
NAME/KEY: mature rHA
LOCATION: 18 to 568
US-09-169-027-11
Query Match 41.7%; Score 43; DB 3; Length 585;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVL 15
Db 426 QRLSGAMDELHNEIL 440

RESULT 26
US-08-453-848-19
Sequence 19, Application US/08453848
Patent No. 5858368
GENERAL INFORMATION:
APPLICANT: Smith, Gale Eugene
APPLICANT: Volvovitz, Franklin
APPLICANT: Wilkinson, Bethanie Bident
APPLICANT: Voznesensky, Andrei I.
APPLICANT: Hackett, Craig Stanway
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,848
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/120,607
FILING DATE: 13-SEPT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: MGS101CIP
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 586 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Harbin/7/94 rHA
FEATURE:
NAME/KEY: HA signal peptide
LOCATION: 1 to 17
FEATURE:
NAME/KEY: mature rHA
LOCATION: 18 to 569
US-08-453-848-19
Query Match 41.7%; Score 43; DB 2; Length 586;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVL 15
Db 427 QRLSGAMDELHNEIL 441

RESULT 27
US-09-169-027-19
Sequence 19, Application US/09169027
Patent No. 6245532
GENERAL INFORMATION:
APPLICANT: Smith, Gale Eugene
APPLICANT: Volvovitz, Franklin
APPLICANT: Wilkinson, Bethanie Bident
APPLICANT: Voznesensky, Andrei I.
APPLICANT: Hackett, Craig Stanway
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/169,027
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/453,848
FILING DATE: 30-MAY-1995
APPLICATION NUMBER: 08/120,607
FILING DATE: 13-SEPT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MGS101CIP

TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Netherlands/13/94 rHA
FEATURE:
NAME/KEY: AcNPV 61K protein signal sequence
LOCATION: 1 to 18
FEATURE:
NAME/KEY: mature rHA
LOCATION: 19 to 571
US-08-453-848-13

Query Match 41.7%; Score 43; DB 2; Length 589;
Best Local Similarity 60.0%; Pred. NO. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ORLYAQBELSNEVL 15
||| | : ||| :
Db 430 ORLSGAMDELHNEIL 444

RESULT 29
US-09-169-027-13
Sequence 13, Application US/09169027
Patent No. 6245532
GENERAL INFORMATION:
APPLICANT: Smith, Gale Eugene
APPLICANT: Volkovitz, Franklin
APPLICANT: Wilkinson, Bethanie Eident
APPLICANT: Voznesensky, Andrei I.
APPLICANT: Hackett, Craig Stanway
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/169,027
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/453,848
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
APPLICATION NUMBER: 08/120,607
FILING DATE: 13-SEPT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MGS101CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795

TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 586 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Harbin/7/94 rHA
FEATURE:
NAME/KEY: HA signal peptide
LOCATION: 1 to 17
FEATURE:
NAME/KEY: mature rHA
LOCATION: 18 to 569
US-09-169-027-19

Query Match 41.7%; Score 43; DB 3; Length 586;
Best Local Similarity 60.0%; Pred. NO. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 ORLYAQBELSNEVL 15
||| | : ||| :
Db 427 ORLSGAMDELHNEIL 441

RESULT 28
US-08-453-848-13
Sequence 13, Application US/08453848
Patent No. 5858368
GENERAL INFORMATION:
APPLICANT: Smith, Gale Eugene
APPLICANT: Volkovitz, Franklin
APPLICANT: Wilkinson, Bethanie Eident
APPLICANT: Voznesensky, Andrei I.
APPLICANT: Hackett, Craig Stanway
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
TITLE OF INVENTION: HEMAGGLUTININ MULTIVALENT VACCINES
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,848
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
APPLICATION NUMBER: 08/120,607
FILING DATE: 13-SEPT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MGS101CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795

INFORMATION FOR SEQ ID NO: 13;
SEQUENCE CHARACTERISTICS:
LENGTH: 592 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Shanghai/4/94 rHA
FEATURE:
NAME/KEY: ACPV 61K protein signal peptide
LOCATION: 1 to 18
FEATURE:
NAME/KEY: mature rHA
LOCATION: 19 to 574
US-08-453-848-17

Query Match 41.7%; Score 43; DB 2; Length 592;
Best Local Similarity 60.0%; Pred. NO. 91;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVL 15
||| :|||
Db 433 QRLGAMDELHNEIL 447

Search completed: May 11, 2004, 13:42:21
Job time : 14.5 secs

INFORMATION FOR SEQ ID NO: 13;
SEQUENCE CHARACTERISTICS:
LENGTH: 589 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Influenza virus
INDIVIDUAL ISOLATE: B/Netherlands/13/94 rHA
FEATURE:
NAME/KEY: ACPV 61K protein signal sequence
LOCATION: 1 to 18
FEATURE:
NAME/KEY: mature rHA
LOCATION: 19 to 571
US-09-169-027-13

Query Match 41.7%; Score 43; DB 3; Length 589;
Best Local Similarity 60.0%; Pred. NO. 90;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVL 15
||| :|||
Db 430 QRLGAMDELHNEIL 444

RESULT 30
US-08-453-848-17
Sequence 17, Application US/08453848
Patent No. 5858368
GENERAL INFORMATION:
APPLICANT: Smith, Gale Eugene
APPLICANT: Volvovitz, Franklin
APPLICANT: Wilkinson, Bethanie Eident
APPLICANT: Voznesensky, Andrei I.
APPLICANT: Hackett, Craig Stanway
TITLE OF INVENTION: A METHOD FOR PRODUCING INFLUENZA
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
STREET: 1201 West Peachtree Street
CITY: Atlanta
STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/453,848
FILING DATE: 30-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/120,607
FILING DATE: 13-SEPT-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MGS101CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)-873-8794
TELEFAX: (404)-873-8795
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:37:37 ; Search time 34.25 Seconds
(without alignments)
162.083 Million cell updates/sec

Title: US-09-171-432A-41

Perfect score: 103

Sequence: 1 QRLKYAQEELSNEVLPPPRK 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications RA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubppaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	103	100.0	20	10 US-09-171-432A-41	Sequence 41, Appl
2	103	100.0	25	10 US-09-171-432A-47	Sequence 47, Appl
3	103	100.0	352	14 US-10-272-459-45	Sequence 45, Appl
4	103	100.0	836	14 US-10-272-459-40	Sequence 40, Appl
5	103	100.0	980	14 US-10-272-459-41	Sequence 41, Appl
6	103	100.0	2227	9 US-09-929-955-12	Sequence 12, Appl
7	103	100.0	2227	13 US-10-104-966-12	Sequence 12, Appl
8	103	100.0	2227	13 US-10-135-988-2	Sequence 2, Appl
9	103	100.0	2227	13 US-10-135-988-6	Sequence 6, Appl
10	98	95.1	2227	13 US-10-135-988-4	Sequence 4, Appl
11	52	50.5	1193	15 US-10-369-493-5713	Sequence 5713, Ap
12	51	49.5	1240	15 US-10-369-493-4031	Sequence 4031, Ap
13	48	46.6	63	12 US-10-424-599-254867	Sequence 254867,
14	48	46.6	442	12 US-10-335-977-5788	Sequence 5788, Ap
15	48	46.6	444	12 US-10-335-977-5789	Sequence 5789, Ap


```

; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match      100.0%; Score 103; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 182 QRLKYAQEELSNEVLPPPRK 201

RESULT 4
US-10-272-459-40
; Sequence 40, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-40

Query Match      100.0%; Score 103; DB 14; Length 836;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 5
US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match      100.0%; Score 103; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;

```

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 6
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match      100.0%; Score 103; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPRK 20
DB 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 7
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: TRIPEP.23AUSCI
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match      100.0%; Score 103; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 8
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175

US-10-135-988-2

Query Match 100.0%; Score 103; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 9
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175

Query Match 100.0%; Score 103; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.6e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 10
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175

Query Match 95.1%; Score 98; DB 13; Length 2227;
Best Local Similarity 95.0%; Pred. No. 2e-05; Mismatches 1; Indels 0; Gaps 0;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLPPPRK 20
|||||
Db 810 QRLKYAQEELSNEVLPPPRK 829

RESULT 11
US-10-369-493-5713
; Sequence 5713, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5713
; LENGTH: 1193
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans

Query Match 50.5%; Score 52; DB 15; Length 1193;
Best Local Similarity 62.5%; Pred. No. 65; Mismatches 3; Indels 0; Gaps 0;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVL 16
|||||
Db 340 QRIKYAREILQKELP 355

RESULT 12
US-10-369-493-4031
; Sequence 4031, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:

APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 4031
LENGTH: 1240
TYPE: PRT
ORGANISM: Neurospora crassa
FEATURE:
NAME/KEY: unsure
LOCATION: (1) .. (1240)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-4031

Query Match 49.5%; Score 51; DB 15; Length 1240;
Best Local Similarity 55.6%; Pred. No. 95;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 QRLKYAQEELNEVLPPPP 18
||| : |||
Db 253 QRSVASEDYENVERSP 270

RESULT 13
US-10-424-599-254867
Sequence 254867, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 254867
LENGTH: 63
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_72168C.1.pep
US-10-424-599-254867

Query Match 46.6%; Score 48; DB 12; Length 63;
Best Local Similarity 52.9%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 3 LKYAQEELNEVLPPPP 19
||| : |||
Db 34 LKYASPELKNIFPIPK 50

RESULT 14
US-10-335-977-5788
Sequence 5788, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 5788:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...442
SEQUENCE DESCRIPTION: SEQ ID NO: 5788:
US-10-335-977-5788

Query Match 46.6%; Score 48; DB 12; Length 442;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 9 ELSNEVLPPPRK 20
| | | | |
Db 414 EADNEELPPPRK 425

RESULT 15
US-10-335-977-5789
Sequence 5789, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 5789:
SEQUENCE CHARACTERISTICS:
LENGTH: 444 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...444
SEQUENCE DESCRIPTION: SEQ ID NO: 5789:
US-10-335-977-5789
Query Match 46.6%; Score 48; DB 12; Length 444;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 ELSNEVLPPPK 20
DB 416 EADNEELPPPK 427
RESULT 16
US-10-424-599-234250
Sequence 234250, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
PRIOR FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 234250
LENGTH: 259
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_53552C.1.pep
US-10-424-599-234250
Query Match 45.6%; Score 47; DB 12; Length 259;
Best Local Similarity 45.5%; Pred. No. 88;
Matches 10; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 1 QRLKYAQEELSN--EVLPPPK 20
DB 106 EMLPYLNEKLKRNVEVPPPK 127
RESULT 17
US-10-425-114-44229
Sequence 44229, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 44229
LENGTH: 355
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700876446_FLI.pep
US-10-425-114-44229
Query Match 45.6%; Score 47; DB 12; Length 355;
Best Local Similarity 45.5%; Pred. No. 96;
Matches 10; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 1 QRLKYAQEELSN--EVLPPPK 20
DB 202 EMLPYLNEKLKRNVEVPPPK 223
RESULT 18
US-10-369-493-1534
Sequence 1534, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 1534
LENGTH: 592
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1534
Query Match 45.6%; Score 47; DB 15; Length 592;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 7 QEELSNEVLPPPK 20
DB 369 EEMENKFAAPPKK 382
RESULT 19
US-10-369-493-5013
Sequence 5013, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 5789:
SEQUENCE CHARACTERISTICS:
LENGTH: 444 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...444
SEQUENCE DESCRIPTION: SEQ ID NO: 5789:
US-10-335-977-5789
Query Match 46.6%; Score 48; DB 12; Length 444;
Best Local Similarity 75.0%; Pred. No. 87;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 9 ELSNEVLPPPK 20
DB 416 EADNEELPPPK 427
RESULT 16
US-10-424-599-234250
Sequence 234250, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
PRIOR FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 234250
LENGTH: 259
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_53552C.1.pep
US-10-424-599-234250
Query Match 45.6%; Score 47; DB 12; Length 259;
Best Local Similarity 45.5%; Pred. No. 88;
Matches 10; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 1 QRLKYAQEELSN--EVLPPPK 20
DB 106 EMLPYLNEKLKRNVEVPPPK 127
RESULT 17
US-10-425-114-44229
Sequence 44229, Application US/10425114
Publication No. US2004003488A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 44229
LENGTH: 355
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 700876446_FLI.pep
US-10-425-114-44229
Query Match 45.6%; Score 47; DB 12; Length 355;
Best Local Similarity 45.5%; Pred. No. 96;
Matches 10; Conservative 4; Mismatches 6; Indels 2; Gaps 1;
QY 1 QRLKYAQEELSN--EVLPPPK 20
DB 202 EMLPYLNEKLKRNVEVPPPK 223
RESULT 18
US-10-369-493-1534
Sequence 1534, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 1534
LENGTH: 592
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1534
Query Match 45.6%; Score 47; DB 15; Length 592;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 7 QEELSNEVLPPPK 20
DB 369 EEMENKFAAPPKK 382
RESULT 19
US-10-369-493-5013
Sequence 5013, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAQEE 9
Db 12 ORLKYAQEE 20

RESULT 21
US-10-424-599-234251
; Sequence 234251, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 234251
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(212)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53553C.1.psp
US-10-424-599-234251

Query Match 43.7%; Score 45; DB 12; Length 212;
Best Local Similarity 40.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 1 ORLKYAQEELSN--EVLPPPRK 20
Db 132 EMVSYLNEXLKRNVEVVPPEKK 153

RESULT 22
US-09-864-761-37656
; Sequence 37656, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5013
; LENGTH: 6642
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5013

Query Match 45.6%; Score 47; DB 15; Length 6642;
Best Local Similarity 45.0%; Pred. No. 2.3e+03;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 ORLKYAQEELSNVLPPEKK 20
Db 1353 RRVSAEEELPKEVIDSRK 1372

RESULT 20
US-09-171-432A-40
; Sequence 40, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khuyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1316
US-09-171-432A-40

Query Match 43.7%; Score 45; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4;

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37656
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC02346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: EST HUMAN HIT: AUI38405.1, EVALUE 7.00e-90
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 1.00e-16
; US-09-864-761-37656

Query Match 43.7%; Score 45; DB 9; Length 263;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 YAEQELNEVLPPPR 19
| : : : | : : |
DB 174 YTEQDLYNHVTPKR 188

RESULT 23
US-10-424-599-234246
; Sequence 234246, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285584
; SEQ ID NO 234246
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53549C.1.pap
; US-10-424-599-234246

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37656
; LENGTH: 263
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC02346.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.2
; OTHER INFORMATION: EST HUMAN HIT: AUI38405.1, EVALUE 7.00e-90
; OTHER INFORMATION: SWISSPROT HIT: P10269, EVALUE 1.00e-16
; US-09-864-761-37656

Query Match 43.7%; Score 45; DB 9; Length 263;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 YAEQELNEVLPPPR 19
| : : : | : : |
DB 174 YTEQDLYNHVTPKR 188

RESULT 23
US-10-424-599-234246
; Sequence 234246, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285584
; SEQ ID NO 234246
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_53549C.1.pap
; US-10-424-599-234246

Query Match 43.7%; Score 45; DB 12; Length 322;
Best Local Similarity 40.9%; Pred. No. 1.7e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 1 QRLKYAOEELSN--EVLPPPRK 20
| : : | : : | : : | : : |
DB 180 EWSVYLNKLEKRNVEVPPPK 201

RESULT 24
US-10-389-566-1970
; Sequence 1970, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1970
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Pisum sativum
; US-10-389-566-1970

Query Match 43.7%; Score 45; DB 16; Length 756;
Best Local Similarity 60.0%; Pred. No. 4.3e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLKYAOEELSNEVL 16
| : : | : : | : : | : : |
DB 64 RLQFAKPNMNEVL 78

RESULT 25
US-10-369-493-3980
; Sequence 3980, Application US/10369493
; Publication No. US2003023675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3980
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1168)
; OTHER INFORMATION: unsure at all Xaa locations
; US-10-369-493-3980

Query Match 43.7%; Score 45; DB 15; Length 1168;
```

Best Local Similarity 43.8%; Pred. No. 6.9e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 YAOEELSNEVLPPPK 20
Db 246 YAEEDSDELLPKPKR 261

RESULT 26

US-09-909-567B-47
; Sequence 47, Application US/0909567B
; Publication No. US2003002257A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto A.
; APPLICANT: Nair, Manoj
; APPLICANT: Chen, Seiyu
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes
; FILE REFERENCE: DEX-0214
; CURRENT APPLICATION NUMBER: US/09/909,567B
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/219,834
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 47
; LENGTH: 1596
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-909-567B-47

Query Match 43.7%; Score 45; DB 10; Length 1596;
Best Local Similarity 43.8%; Pred. No. 9.7e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 3 LKIAQEELSNEVLPPPK 18
Db 799 LNYEEDSDEQVLPPPK 814

RESULT 27

US-10-425-114-57289
; Sequence 57289, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57289
; LENGTH: 235
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE091H01_FLI.pep
US-10-425-114-57289

Query Match 43.2%; Score 44.5; DB 12; Length 235;
Best Local Similarity 55.0%; Pred. No. 1.4e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPPK 20
Db 217 QRLSYAQAPFSG---PPPK 233

RESULT 28

US-10-425-114-72504
; Sequence 72504, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 72504
; LENGTH: 459
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73256H07_FLI.pep
US-10-425-114-72504

Query Match 43.2%; Score 44.5; DB 12; Length 459;
Best Local Similarity 55.0%; Pred. No. 3e+02;
Matches 11; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 QRLKYAQEELSNEVLPPPK 20
Db 441 QRLSYAQAPFSG---PPPK 457

RESULT 29

US-10-424-599-239647
; Sequence 239647, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 239647
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(221)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT MRT3847_58426C.1.pep
US-10-424-599-239647

Query Match 42.7%; Score 44; DB 12; Length 221;
Best Local Similarity 45.0%; Pred. No. 1.6e+02;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 QRLKYAQEELSNEVLPPPK 20
Db 70 RKTQKQEERSNEEPPQE 89

RESULT 30

US-10-369-493-22337
; Sequence 22337, Application US/10369493

; Publication No. US2003033675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 39-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22337
; LENGTH: 303
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22337

Query Match 42.7%; Score 44; DB 15; Length 303;
Best Local Similarity 56.2%; Pred. No. 2.2e+02;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 QRLKYAQEELSNEVLP 16
: ||| : |||
Db 183 RHLKYGPEKNERLP 198

Search completed: May 11, 2004, 13:44:50
Job time : 35.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:23:11 / Search time 47 Seconds
(without alignments)
120.233 Million cell updates/sec

Title: US-09-171-432A-42
Perfect score: 101
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	2	AAW42925 Immunogen
2	101	100.0	21	4	AB669442 Synthetic
3	101	100.0	366	1	AAFP50230 Sequence
4	101	100.0	854	1	AAFP50287 Hepatitis
5	101	100.0	993	1	AAFP50116 Sequence
6	101	100.0	993	1	AAFP50231 Sequence
7	101	100.0	1077	2	AAW95559 A partial
8	101	100.0	1091	2	AAW32426 Translate
9	101	100.0	2227	2	AAW05597 Attenuate
10	101	100.0	2227	2	AAW34074 Hepatitis
11	101	100.0	2227	3	AAW18609 Amino aci
12	101	100.0	2227	3	AAW18607 Amino aci
13	101	100.0	2227	3	AAW18608 Amino aci
14	101	100.0	2227	5	AAW19899 Hepatitis
15	101	100.0	2227	5	ABG31729 Attenuate
16	101	100.0	2227	5	ABG31727 Wild-type
17	101	100.0	2227	5	ABG31728 Hepatitis
18	101	100.0	2227	6	ABU08640 Attenuat
19	101	100.0	2227	6	ABU08641 Attenuat
20	101	100.0	2227	6	ABU08639 Wild type
21	101	100.0	2227	7	ABW00350 Hepatitis
22	96	95.0	2227	1	AAW00066 Sequence
23	84	83.2	939	2	AAW15629 Capsid re
24	65	64.4	25	2	AAW42930 Immunogen
25	65	64.4	26	4	AAW69447 Synthetic

99 41 40.6 108 2 AAW69486 Aaw69486 Hepatitis
100 41 40.6 113 2 AAY48558 Aay48558 Human bre

ALIGNMENTS

RESULT 1
AAW42925
ID AAW42925 standard; peptide; 20 AA.
XX AC AAW42925;
XX DT 28-APR-1998 (first entry)
XX DE Immunogenic Hepatitis A virus peptide YK-1318.
XX KW Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX KW antibody.
XX OS Synthetic.
XX OS Hepatitis A virus.
XX XX W09740147-A1.
XX PD 30-OCT-1997.
XX PF 18-APR-1997; 97WO-US006891.
XX PR 19-APR-1996; 96US-0015644P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
XX DR WPI; 1997-535831/49.
XX PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
XX PT response to HAV in a mammal or to detect the presence of antibodies
XX PT against HAV in a mammal.
XX PS Claim 18; Page 112; 140pp; English.
XX CC Peptides AAW42922-30 are immunogenic peptides corresponding to
XX CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
XX CC substantially similar to a portion of the amino acid sequence of the P2A
XX CC protein of HAV corresponding to amino acids 792-980. The present peptide
XX CC is derived from amino acids 823-842, and has a reactivity of 31.3% with
XX CC acute sera. Compositions containing the peptides can be used to induce an
XX CC immune response to HAV in a mammal. The peptides can also be used to
XX CC detect the presence of antibodies against HAV in mammalian serum. The
XX CC peptides can also be used to make an antibody against HAV by
XX CC administering the peptide to a mammal
XX SQ Sequence 20 AA;

Query Match 100.0%; Score 101; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 1 VLPPPRKMKGLFSQAKISLF 20

RESULT 2
AAB69442
ID AAB69442 standard; peptide; 21 AA.
XX AC AAB69442;
XX DT 20-APR-2001 (first entry)
XX XX

DE Synthetic HAV P2A peptide, SEQ ID NO: 42.
XX Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
KW antigen; major structural capsid polypeptide; HAV antibody detection.
XX Hepatitis A virus.
OS Synthetic.
XX WO200105824-A2.
XX PD 25-JAN-2001.
XX PF 14-JUL-2000; 2000WO-US019267.
XX PR 15-JUL-1999; 99US-0144412P.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Fields HA, Khudyakov YE;
XX DR WPI; 2001-112681/12.
XX PT Synthetic peptides used as antigen sources for enzyme immunoassays
XX PT detecting anti-hepatitis A virus and as vaccines.
XX PS Claim 13; Page 95; 130pp; English.

XX CC The present sequence is one of a number of synthetic peptides which are
XX CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
XX CC comprise antigenic epitopes of the major structural capsid polypeptides
XX CC or non-structural polypeptides of HAV with one or more glutamine
XX CC molecules at the carboxy end of the peptide. The peptides are used to
XX CC detect the presence of antibodies against HAV in mammalian serum, to
XX CC detect the presence of HAV in a human or animal through the binding of
XX CC the peptide to an antibody, to detect acute phase infection by detecting
XX CC IgM antibodies in mammalian serum and detecting convalescence in a
XX CC mammal. The peptides are used to detect or quantify HAV antibodies in
XX CC samples in clinical or research-based assays using immunoblotting,
XX CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
XX CC tracking of radioactive or bioluminescent markers, chromatography or
XX CC electrophoresis. The peptides are used to induce an immune response to
XX CC HAV when administered to a human or animal. Glutamine at the carboxy end
XX CC of the peptides enhances the IgM antibody reactivity

XX SQ Sequence 21 AA;
Query Match 100.0%; Score 101; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.7e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 1 VLPPPRKMKGLFSQAKISLF 20

RESULT 3
AAP50230
ID AAP50230 standard; protein; 366 AA.
XX AC AAP50230;
XX DT 28-NOV-1991 (first entry)
XX DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).
XX KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX KW diagnostic assay.
XX OS Hepatitis A virus.
XX PN EP138704-A.
XX PD 24-APR-1985.

XX 09-OCT-1984; 84EP-00402025.
 XX 14-OCT-1983; 83US-00541836.
 XX 02-MAR-1984; 84US-00585942.
 XX (MERI) MERCK & CO INC.
 XX Hughes JV, Scolnick EM, Tomassini JE;
 XX WPI; 1985-100818/17.
 XX N-PSDB; AAN50274.
 XX New hepatitis A virus surface protein - useful for binding to
 XX neutralising antibodies to the virus.
 XX Claim 21; Page 46-48; 49pp; English.
 XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
 XX surfactant and a reducing agent. The viral proteins are sepd. and the
 XX protein of molecular wt. 33000 daltons is sepd
 XX Sequence 366 AA;
 Query Match 100.0%; Score 101; DB 1; Length 366;
 Best Local Similarity 100.0%; Pred. No. 1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPEKMKGLFSQAKISLF 20
 DB 332 VLPPEKMKGLFSQAKISLF 351
 RESULT 4
 AAP50287
 ID AAP50287 standard; protein; 854 AA.
 XX AAP50287;
 XX 25-MAR-2003 (revised)
 DT 30-NOV-1991 (first entry)
 DE Hepatitis A virus (HAV) peptide corresponding to the capsid protein
 DE region of poliovirus RNA.
 XX Hepatitis A virus assay; antigen; antibody.
 XX Hepatitis A virus.
 XX WO8501517-A.
 XX 11-APR-1985.
 XX 27-SEP-1984; 84WO-US0001552.
 XX 30-SEP-1983; 83US-00537911.
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 XX Racanello VR;
 XX WPI; 1985-098846/16.
 XX N-PSDB; AAN50330.
 XX New hepatitis A virus CDNA - useful in assays for the virus and for
 XX prodn. of the viral antigen and antibodies to it.
 XX Example; Fig 7; 60pp; English.
 XX The inventors claim HAV cDNA and a method for producing it, whereby large
 XX ants. can be obtd. economically. The cDNA is useful in the assay for
 XX detection of HAV quickly and easily and with high sensitivity and

CC specificity. The HAV cDNA is also used in the prodn. of HAV antigen or
 CC antibodies to it. The antibodies may be monoclonal. (Updated on 25-MAR-
 CC 2003 to correct PA field.)
 XX Sequence 854 AA;
 Query Match 100.0%; Score 101; DB 1; Length 854;
 Best Local Similarity 100.0%; Pred. No. 2.3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPEKMKGLFSQAKISLF 20
 DB 823 VLPPEKMKGLFSQAKISLF 842
 RESULT 5
 AAP50116
 ID AAP50116 standard; protein; 993 AA.
 XX AAP50116;
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 30-SEP-1991 (first entry)
 XX Sequence of Hepatitis A virus (HAV) immunogenic peptides VP-1, VP-2, VP-3
 DE and VP-4.
 XX Antigenic protein; immunogen; vaccine.
 XX Hepatitis A virus; (strain CR326).
 XX EPI54587-A.
 XX 11-SEP-1985.
 XX 27-FEB-1985; 85EP-00400369.
 XX 02-MAR-1984; 84US-00585818.
 XX (MERI) MERCK & CO INC.
 XX Linemeyer DL, Menke JG, Rueben RG, Mitra SW;
 XX WPI; 1985-224964/37.
 XX N-PSDB; AAN50139.
 XX New nucleotide sequences coding for hepatitis A virus antigens - useful
 PT for eliciting normal immune response and in vaccines for protecting
 PT against the virus.
 XX Example; Page 11-17; 32pp; English.
 XX Within the sequence in AAN50139 is encoded the information necessary to
 CC make the antigenic proteins of HAV. The sequences encoding for the
 CC structural proteins begin at base 403. The key sub-unit sequences within
 CC VP-1, designated Sequences I, II, III, IV, and V, start, respectively at
 CC 1882, 1963, 1999, 2146, 2347. Other nucleotide sequences which are
 CC valuable as encoding antigenic proteins are the sequences from base 1749
 CC to base 2722; from base 1487 to base 2980 and from base 1644 to base
 CC 2722. The sequence from base 1749 to base 2722 is esp. valuable as a
 CC vector for producing antigen protein. Sequences II-V are claimed. X in
 CC AAP50116 denotes the translation of a stop codon. (Updated on 25-MAR-2003
 CC to correct PI field.) (Updated on 24-OCT-2003 to standardise OS field)
 XX Sequence 993 AA;
 Query Match 100.0%; Score 101; DB 1; Length 993;
 Best Local Similarity 100.0%; Pred. No. 2.7e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPEKMKGLFSQAKISLF 20
 DB 823 VLPPEKMKGLFSQAKISLF 842

Db 959 VLPPPRKMKGLFSQAKISLF 978

RESULT 6
AAP50231
ID AAP50231 standard; protein; 993 AA.
AC AAP50231;
XX
XX 28-NOV-1991 (first entry)
DT
XX
DE Sequence encoded by partial sequence of hepatitis A virus (HAV),
DE including surface protein (VP-1).
XX
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
KW diagnostic assay.
XX
XX Hepatitis A virus.
OS
XX
XX Key Location/Qualifiers
FH 628..993
FT Protein
FT /note= "claimed; X denotes translated stop codons and
FT unspecified triplets"
XX
XX EPI38704-A.
PN
XX
XX 24-APR-1985.
PD
XX
XX 09-OCT-1984; 84EP-00402025.
PF
XX 14-OCT-1983; 83US-00541836.
PR
PR 02-MAR-1984; 84US-00585942.
XX
XX (MERI) MERCK & CO INC.
PA
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
PI
XX
XX WPI; 1985-100818/17.
DR
DR N-PSDB; AAN50274.
XX
XX New hepatitis A virus surface protein - useful for binding to
FT neutralising antibodies to the virus.
PT
XX
XX Disclosure; Page 17-23; 49pp; English.
PS
XX
XX VP1 is isolated by solubilisation of the intact virus in an aq. anionic
CC surfactant and a reducing agent. The viral proteins are sepd. and the
CC protein of molecular wt. 33000 daltons is sepd
CC
XX
XX Sequence 993 AA;
SQ

Query Match 100.0%; Score 101; DB 1; Length 993;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 959 VLPPPRKMKGLFSQAKISLF 978

RESULT 7
AAW95559
ID AAW95559 standard; protein; 1077 AA.
XX
XX
XX AAW95559;
AC
DT 28-APR-1999 (first entry)
XX
XX A partial hepatitis A virus (HAV) protein.
DE
XX
XX Hepatitis A virus protein; HAV; P2 region;
KW cell-culture-adapted HAV strain; infection; accelerated growth.
XX

OS Hepatitis A virus.
XX
XX US5849562-A.
XX
PD 15-DEC-1998.
XX
XX 06-JUN-1995; 95US-00468926.
XX
XX 30-SEP-1983; 83US-00537911.
PR 27-SEP-1984; 84US-0054942.
PR 06-OCT-1988; 88US-00256135.
PR 06-NOV-1991; 91US-00788262.
XX
XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
PA
XX Emerson SU, Purcell RH;
PI
XX WPI; 1999-094412/08.
DR N-PSDB; AAX01006.
XX
XX Chimeric hepatitis A virus strains - with P2 region from cell-culture-
FT adapted strain in wild-type genome.
XX
XX Disclosure; Fig 7A-L; 36pp; English.
XX
XX The present sequence represents a partial hepatitis A virus (HAV)
CC protein. The specification describes a DNA construct consisting of a wild
CC -type HAV genome in which the P2 region is replaced by the P2 region from
CC a cell-culture-adapted HAV strain. The construct is used to demonstrate
CC that mutations in the P2 region of a cell-culture-adapted HAV strain are
CC sufficient for establishment of infection and accelerated growth in cell
CC culture
XX
XX Sequence 1077 AA;
SQ

Query Match 100.0%; Score 101; DB 2; Length 1077;
Best Local Similarity 100.0%; Pred. No. 3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 1046 VLPPPRKMKGLFSQAKISLF 1065

RESULT 8
AAR32426
ID AAR32426 standard; protein; 1091 AA.
XX
XX AAR32426;
AC
XX 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 17-DEC-2001 (revised)
DT 10-JUN-1993 (first entry)
XX
XX Translated from 5' region of Hepatitis A Virus genomic clone.
DE HAV HM-175; chronic liver disease; picornavirus.
XX
XX Hepatitis A virus.
OS
XX
XX Key Location/Qualifiers
FH 1..711
FT /note= "X's correspond to nonsense codons, i.e. this
FT region is not an ORF"
FT 238..1091
FT /label= ORF
FT /note= "second putative initiation codon at position 240"
XX
XX USN7789262-N.
PN
XX
XX 15-DEC-1992.
PD
XX

PF 06-NOV-1991; 91US-00788262.
 XX 30-SEP-1983; 88US-00536911.
 PR 27-SEP-1984; 84US-00854942.
 PR 06-OCT-1988; 88US-00256135.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICE.
 XX
 XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
 PI Racanello VR, Baroudy BM, Emerson SU;
 XX
 XX WPI; 1993-067429/08.
 DR N-PSDB; AAQ36934.
 DR
 XX
 XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn. of
 PT antigen and antibodies.
 PT
 XX
 PS Disclosure; Fig 7; 65pp; English.
 XX
 XX HAV virion RNA was extracted from the livers of marmosets which had been
 CC inoculated with HAV (the HAV had previously been passaged twice in
 CC marmosets). The RNA was used to prepare ds cDNA clones by standard
 CC methods. Clones contg. inserts which hybridised to RNA from HAV-infected
 CC African Green Monkey Kidney cells were selected for further analysis. A
 CC 7.4kb restriction map (about 99% of the HAV genome) was constructed from
 CC 5' overlapping inserts. The sequence of the first 3.3kb (approx.) from the
 CC 5'-terminus was determined. An amino acid sequence was decoded from the
 CC entire clone and an open reading frame was identified starting at
 CC position 238. A comparison of the predicted HAV amino acid sequences with
 CC the known capsid protein sequences of other picornaviruses (poliovirus,
 CC foot and mouth disease virus and encephalomyelitis virus) revealed areas
 CC of local homology. (Note: Revised entry submitted to correct the patent
 CC number format of US Government-owned NTIS applications to prevent clashes
 CC with ongoing US granted patent numbers. For further information please
 CC visit the Derwent web site at www.derwent.com/dwpi/updates/ntis.us.html.)
 CC (Updated on 25-MAR-2003 to correct PI field.) (Updated on 27-AUG-2003 to
 CC correct OS field.)
 XX
 XX
 SQ Sequence 1091 AA;
 Query Match 100.0%; Score 101; DB 2; Length 1091;
 Best Local Similarity 100.0%; Pred. No. 3e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 DB 1060 VLPPPRKMKGLFSQAKISLF 1079
 RESULT 9
 AAR05697
 ID AAR05697 standard; protein; 2227 AA.
 XX
 AC AAR05697;
 XX
 XX 24-OCT-2003 (revised)
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 15-AUG-1990 (first entry)
 XX
 XX Attenuated hepatitis A virus.
 XX
 XX Hepatitis A virus; vaccine; attenuated.
 KW
 XX
 XX Hepatitis A virus; strain HM-175.
 OS
 XX
 FH Key Location/Qualifiers
 FT Region 1..23
 FT /label= VP4 = 1A
 FT Region 24..245
 FT /label= VP2 = 1B
 FT Region 246..491
 FT /label= VP3 = 1C

FT Region 492..791
 FT /label= VP1 = 1D
 FT Region 792..980
 FT /label= 2A
 FT Region 981..1087
 FT /label= 2B
 FT Region 1088..1422
 FT /label= 2C
 FT Region 1423..1496
 FT /label= 3A
 FT Region 1497..1519
 FT /label= 3B = VPg
 FT Region 1520..1738
 FT /label= 3C
 FT Region 1739..2227
 FT /label= 3D
 XX
 PN US4894228-A.
 XX
 XX 16-JAN-1990.
 PD
 XX 12-JUL-1988; 88US-00217824.
 PF
 XX 19-SEP-1984; 84US-00652067.
 PR
 XX 09-SEP-1986; 86US-00905146.
 PR
 XX (USSH) US DEPT HEALTH & HUMAN SERVICE.
 PA
 XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;
 PI Daemer RJ, Gust ID;
 XX
 XX WPI; 1990-075557/10.
 DR N-PSDB; AAQ03512.
 DR
 XX Vaccine against hepatitis A virus infection - comprises novel attenuated
 PT hepatitis A virus strain.
 PT
 XX Claim 1; Fig 1; 18pp; English.
 PS
 XX
 CC The attenuated HAV is useful for inducing protective immunity against
 CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
 CC several nucleotide changes distributed throughout the genome, is
 CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
 CC suitable for use as an HAV vaccine. It is noted that not all the changes
 CC are necessary for attenuation and use as a vaccine. (Updated on 25-MAR-
 CC 2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC (Updated on 27-AUG-2003 to correct OS field.) (Updated on 24-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 2; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 DB 823 VLPPPRKMKGLFSQAKISLF 842
 RESULT 10
 AAW34074
 ID AAW34074 standard; protein; 2227 AA.
 XX
 AC AAW34074;
 XX
 XX 17-OCT-2003 (revised)
 DT 27-OCT-1998 (first entry)
 DT
 XX Hepatitis A virus HM-175 protein sequence.
 DE
 XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection;
 KW vaccine.

XX OS Hepatitis A virus; HM-175.

XX FH Key Location/Qualifiers

FT Protein 1. .23

FT Protein /label= VP4

FT Protein 24. .245

FT Protein /label= VP2

FT Protein 246. .491

FT Protein /label= VP3

FT Protein 492. .791

FT Protein /label= VP1

FT Protein 792. .980

FT Protein /label= 2A

FT Protein 981. .1087

FT Protein /label= 2B

FT Protein 1088. .1422

FT Protein /label= 2C

FT Protein 1423. .1496

FT Protein /label= 3A

FT Protein 1497. .1519

FT Protein /label= 3B

FT Protein 1520. .1738

FT Protein /label= 3C

FT Protein 1739. .2227

FT Protein /label= 3D

XX WO9740166-A2.

XX 30-OCT-1997.

XX 18-APR-1997; 97WO-US006506.

XX 19-APR-1996; 96US-0015642P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Raychaudhuri G, Emerson SU, Purcell RH;

XX WPI; 1997-535850/49.

XX N-PSDB; AAT93023.

XX Human attenuated HAV genome containing simian HAV 2C gene - useful as

XX vaccines against HAV infection.

XX Disclosure; Fig 13A-D; 66pp; English.

XX This protein sequence is encoded by the human hepatitis A virus (HAV) HM-

XX 175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained

XX by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA

XX construct (I) comprises a genome of HAV, where the genome is a human

XX attenuated HAV genome in which a region of the 2C gene has been replaced

XX by a corresponding region from a 2C gene of a simian AGM-27 HAV genome

XX (see AAT93024). The region of the 2C gene from AGM-27 contained in the

XX construct preferably encodes amino acids 120-328 of the 2C protein, amino

XX acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript

XX of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3)

XX a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host

XX cell containing the HAV of (3). (1) or its RNA transcript, can be used as

XX a vaccine for preventing HAV in a mammal. (1) or the RNA transcript can

XX also be used to stimulate the production of protective antibodies in the

XX mammal. (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 2227 AA;

XX Query Match 100.0%; Score 101; DB 2; Length 2227;

XX Best Local Similarity 100.0%; Pred. No. 6.1e-07;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

DB 823 VLPPPRKMKGLFSQAKISLF 842

OS Hepatitis A virus.

XX US6113912-A.

XX PN

XX RESULT 11

XX AAB18609 standard; protein; 2227 AA.

XX AC AAB18609;

XX DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX HAV 4380.

XX Hepatitis A virus.

XX US6113912-A.

XX PD 05-SEP-2000.

XX PF 07-JUN-1995; 95US-00475886.

XX PR 18-SEP-1992; 92US-00947338.

XX PR 17-SEP-1993; 93WO-US008610.

XX PR 17-APR-1995; 95US-00397232.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PA D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

XX N-PSDB; AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

XX line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type.

XX Disclosure; Col 93-104; 72pp; English.

XX The present sequence is derived from a live attenuated hepatitis A virus

XX (HAV) of the invention, designated HAV 4380. The sequence is produced by

XX modifying wild type HAV strain HM-174. The HAV of the invention are

XX adapted to growth in the human fibroblast-like cell line MRC-5. The HAV

XX is able to propagate in MRC-5 cells and retain appropriate attenuation.

XX It is useful as a live vaccine for prophylaxis of hepatitis A in humans

XX and other primates

XX Sequence 2227 AA;

XX Query Match 100.0%; Score 101; DB 3; Length 2227;

XX Best Local Similarity 100.0%; Pred. No. 6.1e-07;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

DB 823 VLPPPRKMKGLFSQAKISLF 842

XX RESULT 12

XX AAB18607

XX ID AAB18607 standard; protein; 2227 AA.

XX AC AAB18607;

XX DT 15-JAN-2001 (first entry)

XX DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

XX US6113912-A.

XX PN

XX 05-SEP-2000.
XX
XX
XX PF 07-JUN-1995; 95US-00475886.
XX
XX PR 18-SEP-1992; 92US-00947338.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI; 2000-586464/55.
XX DR N-PSDB; AAA75476.
XX
XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX PT infection, has modified genome compared to wild type.
XX
XX PS Disclosure; Fig 6A-K; 72pp; English.
XX
XX CC The present sequence is derived from a wild type hepatitis A virus (HAV)
XX strain HM-174. The sequence is modified to produce HAV which are adapted
XX CC to growth in the human fibroblast-like cell line MRC-5. The HAV is able
XX CC to propagate in MRC-5 cells and retain appropriate attenuation. It is
XX CC useful as a live vaccine for prophylaxis of hepatitis A in humans and
XX CC other primates
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPrKMKGLFSQAKISLF 20
DB 823 VLPPrKMKGLFSQAKISLF 842
RESULT 13
AAB18608
ID AAB18608 standard; protein; 2227 AA.
XX
XX AC AAB18608;
XX
XX DT 15-JAN-2001 (first entry)
XX
XX DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.
XX
XX KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;
XX P-35 virus.
XX
XX OS Hepatitis A virus.
XX
XX PN US6113912-A.
XX
XX PD 05-SEP-2000.
XX
XX PF 07-JUN-1995; 95US-00475886.
XX
XX PR 18-SEP-1992; 92US-00947338.
XX PR 17-SEP-1993; 93WO-US008610.
XX PR 17-APR-1995; 95US-00397232.
XX
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
XX PI D'hondt E, Purcell RH, Emerson SU, Funkhouser AW;
XX
XX WPI; 2000-586464/55.
XX DR N-PSDB; AAA75477.
XX
XX PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type.
XX
XX PS Disclosure; Col 67-78; 72pp; English.
XX
XX CC The present sequence is derived from passage 35 of a wild type hepatitis
XX A virus (HAV) strain HM-174. The resulting virus is designated P-35
XX CC virus. The sequence is modified to produce HAV which are adapted to
XX CC growth in the human fibroblast-like cell line MRC-5. The HAV is able to
XX CC propagate in MRC-5 cells and retain appropriate attenuation. It is useful
XX CC as a live vaccine for prophylaxis of hepatitis A in humans and other
XX CC primates
XX
XX SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPrKMKGLFSQAKISLF 20
DB 823 VLPPrKMKGLFSQAKISLF 842
RESULT 14
AAE19899
ID AAE19899 standard; protein; 2227 AA.
XX
XX AC AAE19899;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX DE Hepatitis A virus (HAV) protein.
XX
XX KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX OS Hepatitis A virus.
XX
XX PN WO200213855-A2.
XX
XX PD 21-FEB-2002.
XX
XX PF 15-AUG-2001; 2001WO-1B001808.
XX
XX PR 17-AUG-2000; 2000US-0225767P.
XX PR 29-AUG-2000; 2000US-0229175P.
XX PR 03-NOV-2000; 2000US-00705547.
XX
XX PA (TRIP-) TRIPEP AB.
XX
XX PI Sallberg M, Hultgren C;
XX
XX DR WPI; 2002-241837/29.
XX DR N-PSDB; AAD31766.
XX
XX PT Vaccine compositions for treating and preventing disease, preferably
XX PT hepatitis C virus infection, comprises ribavirin and antigen that has
XX PT epitope present in hepatitis C virus.
XX
XX PS Claim 11; Page 82-87; 120pp; English.
XX
XX CC The invention relates to a composition comprising ribavirin and an
XX CC antigen preferably non structural 3 protein (NS3)/4A fragment of
XX CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
XX CC sequence. The composition is useful for enhancing an immune response to a
XX CC hepatitis C antigen in humans, domestic, sport or pet species and as
XX CC vaccines for treating and preventing HCV infections. The composition is
XX CC also useful for treating viral, bacterial, fungal diseases and cancer.
XX
XX SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
 |||||
 DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 15
 ABG31729
 ID ABG31729 standard; protein; 2227 AA.
 XX AC ABG31729;
 XX DT 29-AUG-2003 (revised)
 XX DT 29-NOV-2002 (first entry)
 XX DE Attenuated Hepatitis A virus (HAV) 4390 polypeptide.
 XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
 XX KW HAV 4380.
 XX OS Hepatitis A virus; strain HM-175.
 XX DN US6423318-B1.
 XX PD 23-JUL-2002.
 XX PF 31-AUG-2000; 2000US-00653499.
 XX PR 17-SEP-1993; 93WO-US008610.
 XX PR 17-APR-1995; 95US-00397232.
 XX PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 XX DR WPI; 2002-680946/73.
 XX DR N-PSDB; ABS52789.
 XX PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 XX PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 XX PS Disclosure; Col 93-104; 71pp; English.
 XX CC The invention relates to a polynucleotide which encodes a hepatitis A
 XX CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 XX CC line). The polynucleotide is useful for preparing a vaccine against
 XX CC hepatitis A virus infection. This sequence represents an attenuated
 XX CC hepatitis A virus 4830 polypeptide. (Updated on 29-AUG-2003 to
 XX CC standardise OS field)

QY 1 VLPPPRKMKGLFSQAKISLF 20
 |||||
 DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 16
 ABG31727
 ID ABG31727 standard; protein; 2227 AA.
 XX AC ABG31727;
 XX XX 29-AUG-2003 (revised)

DT 29-NOV-2002 (first entry)
 XX DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
 XX KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
 XX OS Hepatitis A virus; strain HM-175.
 XX PN US6423318-B1.
 XX PD 23-JUL-2002.
 XX PF 31-AUG-2000; 2000US-00653499.
 XX PR 17-SEP-1993; 93WO-US008610.
 XX PR 17-APR-1995; 95US-00397232.
 XX PR 07-JUN-1995; 95US-00475886.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 XX DR WPI; 2002-680946/73.
 XX DR N-PSDB; ABS52789.
 XX PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 XX PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 XX PS Disclosure; Fig 6; 71pp; English.
 XX CC The invention relates to a polynucleotide which encodes a hepatitis A
 XX CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 XX CC line). The polynucleotide is useful for preparing a vaccine against
 XX CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 XX CC strain HM-175 polypeptide. (Updated on 29-AUG-2003 to standardise OS
 XX CC field)

QY 1 VLPPPRKMKGLFSQAKISLF 20
 |||||
 DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 17
 ABG31728
 ID ABG31728 standard; protein; 2227 AA.
 XX AC ABG31728;
 XX DT 29-NOV-2002 (first entry)
 XX DE Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.
 XX KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
 XX KW virucide; mutant; pHAV/7; mutein.
 XX OS Hepatitis A virus; strain HM-175.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 XX FT Misc-difference 764
 XX FT Misc-difference 821 /note= "Wild-type Glu substituted by Val"
 XX FT Misc-difference 821 /note= "Wild-type Asn substituted by Ser"
 XX FT Misc-difference 963 /label= Wild-type Lys substituted by Arg
 XX FT Misc-difference 1052

FT /note= "Wild-type Ala substituted by Val"
 FT Misc-difference 1062
 FT /note= "Wild-type Gly substituted by Ala"
 FT Misc-difference 1118
 FT /note= "Wild-type Lys substituted by Met"
 FT Misc-difference 1151
 FT /note= "Wild-type Glu substituted by Lys"
 FT Misc-difference 1163
 FT /note= "Wild-type Phe substituted by Ser"
 FT Misc-difference 1277
 FT /note= "Wild-type Val substituted by Ile"
 FT Misc-difference 1500
 FT /note= "Wild-type His substituted by Tyr"
 FT Misc-difference 1805
 FT /note= "Wild-type Asp substituted by Asn"
 FT Misc-difference 1930
 FT /note= "Wild-type Ser substituted by Thr"
 XX
 PW US6423318-B1.
 XX
 PD 23-JUL-2002.
 XX
 PF 31-AUG-2000; 2000US-00653499.
 XX
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 XX
 DR WPI; 2002-680946/73.
 DR N-PSDB; ABS52788.
 XX
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection.
 XX
 PS Example 3; Col 67-78; 71pp; English.
 XX
 CC The invention relates to a polynucleotide which encodes a hepatitis A
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
 CC line). The polynucleotide is useful for preparing a vaccine against
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus
 CC mutant strain HM-175/7 (pHAV/7) polypeptide
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 5; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPFRKMKGLFSQAKISLF 20
 |||||
 DB 823 VLPFRKMKGLFSQAKISLF 842
 RESULT 18
 ABU08640
 ID ABU08640 standard; protein; 2227 AA.
 XX
 AC ABU08640;
 XX
 DT 23-OCT-2003 (revised)
 DT 03-JUN-2003 (first entry)
 XX
 DE Attenuated (pass35) hepatitis A virus strain HM-175.
 XX
 KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
 KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
 XX
 OS Hepatitis A virus; strain HM-175.

XX US2002176869-A1.
 PN
 XX 28-NOV-2002.
 PD
 XX 29-APR-2002; 2002US-00135988.
 PF
 XX 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 PR 31-AUG-2000; 2000US-00653499.
 XX
 PA (FUNK/) FUNKHOUSER A W.
 PA (EMER/) EMERSON S U.
 PA (PURC/) PURCELL R H.
 PA (DHON/) D'HONDT E.
 XX
 PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
 XX
 DR WPI; 2003-352605/02.
 DR N-PSDB; ABX93474.
 XX
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
 PT useful in vaccines for protecting primates against hepatitis infection
 PT and disease.
 XX
 PS Example 3; Fig 6; 70pp; English.
 XX
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
 CC a vaccine for protecting primates against hepatitis infection and
 CC disease. This is the amino acid sequence of an attenuated (pass 35)
 CC human hepatitis A virus strain HM-175. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 2227 AA;
 Query Match 100.0%; Score 101; DB 6; Length 2227;
 Best Local Similarity 100.0%; Pred. No. 6.1e-07;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPFRKMKGLFSQAKISLF 20
 |||||
 DB 823 VLPFRKMKGLFSQAKISLF 842
 RESULT 19
 ABU08641
 ID ABU08641 standard; protein; 2227 AA.
 XX
 AC ABU08641;
 XX
 DT 23-OCT-2003 (revised)
 DT 03-JUN-2003 (first entry)
 XX
 DE Attenuated hepatitis A virus (4380) strain HM-175.
 XX
 KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
 KW vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.
 XX
 OS Hepatitis A virus; strain HM-175.
 XX
 PN US2002176869-A1.
 PD
 XX 28-NOV-2002.
 XX
 PF 29-APR-2002; 2002US-00135988.
 XX
 PR 18-SEP-1992; 92US-00947338.
 PR 17-SEP-1993; 93WO-US008610.
 PR 17-APR-1995; 95US-00397232.
 PR 07-JUN-1995; 95US-00475886.
 PR

PR 31-AUG-2000; 2000US-00653499.
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
PI Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93475.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
XX Disclosure; Page 45-51; 70pp; English.
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of an attenuated human
CC hepatitis A virus (4380) strain HM-175. (Updated on 23-OCT-2003 to
CC standardise OS field)
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842
RESULT 20
ABU08639
ID ABU08639 standard; protein; 2227 AA.
XX
XX ABU08639;
XX
DT 23-OCT-2003 (revised)
DT 03-JUN-2003 (first entry)
XX
XX Wild type human hepatitis A virus strain HM-175.
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
XX vaccine; MRC-5 cell; hepatitis infection.
XX Hepatitis A virus; strain HM-175.
XX
XX US2002176869-A1.
XX
XX 28-NOV-2002.
XX
XX 29-APR-2002; 2002US-00135988.
XX
XX 18-SEP-1992; 92US-00947338.
XX 17-SEP-1993; 93WO-US008610.
XX 17-APR-1995; 95US-00397232.
XX 07-JUN-1995; 95US-00475886.
XX 31-AUG-2000; 2000US-00653499.
XX
XX (FUNK/) FUNKHOUSER A W.
PA (EMER/) EMERSON S U.
PA (PURC/) PURCELL R H.
PA (DHON/) D'HONDT E.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'hondt E;
XX
XX WPI; 2003-352605/02.
DR N-PSDB; ABX93473.

XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
PT useful in vaccines for protecting primates against hepatitis infection
PT and disease.
XX
XX Disclosure; Fig 6; 70pp; English.
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
CC a vaccine for protecting primates against hepatitis infection and
CC disease. This is the amino acid sequence of wild type human hepatitis A
CC virus strain HM-175. (Updated on 23-OCT-2003 to standardise OS field)
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 6; Length 2227;
Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842
RESULT 21
ABW00350
ID ABW00350 standard; protein; 2227 AA.
XX
XX ABW00350;
XX
DT 15-JAN-2004 (first entry)
XX
XX Hepatitis A virus protein.
XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
XX virucide.
XX
XX Hepatitis A virus.
XX
XX US2002136740-A1.
XX
XX 26-SEP-2002.
XX
XX 15-AUG-2001; 2001US-00929955.
XX
XX 17-AUG-2000; 2000US-0225767P.
XX 29-AUG-2000; 2000US-0229175P.
XX
XX (SALL/) SALLBERG M.
XX (HULT/) HULTGREN C.
XX
XX Sallberg M, Hultgren C;
XX
XX WPI; 2003-764978/72.
XX N-PSDB; AAD60867.
XX
XX Vaccine compositions for treating and preventing disease, preferably
PT hepatitis C virus infection, comprises ribavirin and antigen that has
PT epitope present in hepatitis C virus.
XX
XX Claim 11; Page 45-51; Opp; English.
XX
XX The invention relates to a composition comprising ribavirin and an
CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
CC is useful in enhancing the immune response to a hepatitis C antigen where
CC the composition is delivered to an animal identified as requiring an
CC enhanced immune response. The vaccine is useful in the treatment and
CC prevention of hepatitis C infection. The present sequence is Hepatitis A
XX virus protein
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 7; Length 2227;

Best Local Similarity 100.0%; Pred. No. 6.1e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPFPRKMKGLFSQAISLF 20
Db 823 VLPFPRKMKGLFSQAISLF 842

RESULT 22
AAP60066
ID AAP60066 standard; protein; 2227 AA.
XX AC AAP60066;
XX DT 25-MAR-2003 (revised)
XX DT 26-JUN-1991 (first entry)
XX DE Sequence of viral I434 polypeptide encoded by the complete nucleotide
DE sequence of the HAV genome.
XX DE
XX DE
XX KW Diagnosis; vaccine; passive immunotherapy.
XX OS Hepatitis A virus.

XX Key Location/Qualifiers
XX FFH 1..245
FT Region /label= P1.1A
FT FT 246..491
FT Region /label= 1B
FT FT 492..836
FT Region /label= 1C
FT FT 837..980
FT Region /label= P2.2A
FT FT 981..1076
FT FT /label= 2B
FT Region 1077..1142
FT FT /label= 2C
FT Region 1423..1484
FT FT /label= P3.3A
FT Region 1485..1507
FT FT /label= 3B
FT Region 1508..1678
FT FT /label= 3C
FT Region 1679..2227
FT FT /label= 3D
XX PN EP199480-A.
XX PD 29-OCT-1986.
XX PF 03-APR-1986; 86EP-00302465.
XX PR 03-APR-1985; 85US-00719329.
XX (CHIR) CHIRON CORP.
XX PI Dina D, Potter SJ, Vannest GA, Caput D;
XX DR WPI; 1986-286213/44.
XX DR N-PSDB; AAN60080.
XX PT Hepatitis A virus nucleotide sequence and polypeptide - and use in prodn.
XX ST of vaccines and diagnostic probes.
XX PS Claim 5; Fig 1; 18pp; English.
XX
XX AAN60080 and oligonucleotide fragments are useful in detection of
CC hepatitis A virus; transformed hosts may be used for expression of
CC polypeptides and fragments useful in vaccines without risk of infection
CC by the virus or in prodn. of particles which are capable of inducing
CC immunocompetent B cells for passive immunotherapy. Pref. epitope is
CC derived from AAs 445-657 or 792-848 of the HAV polypeptide sequence
CC (AAP60066). (Updated on 25-MAR-2003 to correct PA field.)
CC

CC cyno-HAV. The other two cleavage sites are the same. Two residues have
CC been identified as part of the immuno- dominant region (see feature
CC table) and are different to those in the same position in human HAV. The
CC protein and peptides derived from it can be used in the prepn. of
CC vaccines for the prevention of HAV infection. See also AAR15056. (Note:
CC Revised entry submitted to correct the patent number format of US
CC Government-owned NTIS applications to prevent clashes with ongoing US
CC granted patent numbers. For further information please visit the Derwent
CC web site at www.derwent.com/dwpi/updates/ntis-us.html.) (Updated on 25-
CC MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS
CC field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
SQ Sequence 839 AA;

Query Match 83.2%; Score 84; DB 2; Length 839;
Best Local Similarity 83.3%; Pred. No. 0.00012;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKIS 18
:|||||:|||||:|||||
DB 822 ILPPPRKMKGLFSQSKIS 839

RESULT 24
AAW42930
ID AAW42930 standard; peptide; 25 AA.

XX AC AAW42930;
XX DT 28-APR-1998 (first entry)
XX XX Immunogenic Hepatitis A virus peptide YK-1665.

XX Immunogenic peptide; immunogenic epitope; P2A protein; immune response;
XX antibody.

XX Synthetic.
XX OS Hepatitis A virus.

XX PN WO9740147-A1.

XX PD 30-OCT-1997.

XX PF 18-APR-1997; 97WO-US006891.

XX PR 19-APR-1996; 96US-0015644P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 1997-535831/49.

XX Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an immune
PT response to HAV in a mammal or to detect the presence of antibodies
PT against HAV in a mammal.

XX PS Claim 18; Page 112; 140pp; English.

XX Peptides AAW42922-30 are immunogenic peptides corresponding to
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are
CC substantially similar to a portion of the amino acid sequence of the P2A
CC protein of HAV corresponding to amino acids 792-980. Compositions
CC containing the peptides can be used to induce an immune response to HAV
CC in a mammal. The peptides can also be used to detect the presence of
CC antibodies against HAV in mammalian serum. The peptides can also be used
CC to make an antibody against HAV by administering the peptide to a mammal
XX

SQ Sequence 25 AA;

Query Match 64.4%; Score 65; DB 2; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLF 12
:|||||:|||||
DB 14 VLPPPRKMKGLF 25

RESULT 25

AA69447

ID AA69447 standard; peptide; 26 AA.

XX AC AA69447;

XX DT 20-APR-2001 (first entry)

XX DE Synthetic HAV P2A peptide, SEQ ID NO: 47.

XX KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;
XX antigen; major structural capsid polypeptide; HAV antibody detection.

XX OS Hepatitis A virus.

XX OS Synthetic.

XX PN WO200105824-A2.

XX PD 25-JAN-2001.

XX PF 14-JUL-2000; 2000WO-US019267.

XX PR 15-JUL-1999; 99US-0144412P.

XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX PI Fields HA, Khudyakov YE;

XX DR WPI; 2001-112681/12.

XX Synthetic peptides used as antigen sources for enzyme immunoassays
PT detecting anti-hepatitis A virus and as vaccines.

XX PS Claim 13; Page 98; 130pp; English.

XX The present sequence is one of a number of synthetic peptides which are
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides
CC comprise antigenic epitopes of the major structural capsid polypeptides
CC or non-structural polypeptides of HAV with one or more glutamine
CC molecules at the carboxy end of the peptide. The peptides are used to
CC detect the presence of antibodies against HAV in mammalian serum, to
CC detect the presence of an antibody, to detect acute phase infection by detecting
CC the peptide to an antibody, to detect acute phase convalescence in a
CC IGM antibodies in mammalian serum and detecting convalescence in a
CC mammal. The peptides are used to detect or quantify HAV antibodies in
CC samples in clinical or research-based assays using immunoblotting,
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,
CC tracking of radioactive or bioluminescent markers, chromatography or
CC electrophoresis. The peptides are used to induce an immune response to
CC HAV when administered to a human or animal. Glutamine at the carboxy end
CC of the peptides enhances the IGM antibody reactivity

XX SQ Sequence 26 AA;

Query Match 64.4%; Score 65; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0038;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLF 12
:|||||:|||||
DB 14 VLPPPRKMKGLF 25

RESULT 26

AAR86001

ID AAR86001 standard; protein; 1025 AA.

XX

AC AAR86001;
 XX 04-JUN-1996 (first entry)
 DT Human dihydropyrimidine dehydrogenase.
 DE Bovine; liver; human; dihydropyrimidine dehydrogenase; DPD;
 XX pyrimidine catabolism; 5,6-dihydropyrimidine; pyrimidine analogue;
 KW fluoropyrimidine; anticancer drug; 5-fluorouracil; FURA; cancer;
 KW frameshift mutation.
 XX Homo sapiens.
 OS W09528489-A1.
 XX W09528489-A1.
 XX 26-OCT-1995.
 PD 13-APR-1995; 95WO-US004567.
 XX 13-APR-1994; 94US-00227357.
 PR (UABR-) UAB RES FOUND.
 XX Diasio RB, Lu Z, Zhang R, Johnson M, Cheng X;
 PI WPI; 1995-373803/48.
 DR N-PSDB; AAT03133.
 XX Novel dihydro:pyrimidine dehydrogenase gene - used to optimise 5-
 PT fluoro:uracil doses given to cancer patients.
 PS Claim 9; Page 140-56; 207pp; English.
 XX This sequence represents human liver dihydropyrimidine dehydrogenase
 CC (DPD). DPD catalyses the initial and rate limiting step in pyrimidine
 CC catabolism, the reduction of pyrimidines to 5,6-dihydropyrimidines. DPD
 CC is a complex enzyme consisting of two identical subunits, containing FMN,
 CC FAD and iron-sulphur centers, and utilising NADPH as a cofactor. DPD has
 CC also been shown to catalyse the reduction of various pyrimidine analogues
 CC including the fluoropyrimidine anticancer drug 5-fluorouracil (FURA). Up
 CC to 85% of administered FURA may be catabolised by DPD, and it therefore
 CC governs the effectiveness of FURA as an anticancer drug. DPD genes or
 CC fragments of them may be used in the detection of DPD in a sample, esp.
 CC isolated from a cancer patient. According to the amount of DPD detected,
 CC a therapeutically effective amount of FURA may be determined and
 CC administered. DPD deficiency, leading to life-threatening toxicity on
 CC exposure to FURA, in a human caused by a frameshift mutation may be
 CC determined by means of a molecular biological assay to detect the
 CC deletion of an A residue at codon 318 within the DPD-coding region
 XX SQ Sequence 1025 AA;
 Query Match 52.0%; Score 52.5; DB 2; Length 1025;
 Best Local Similarity 63.2%; Pred. No. 15;
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 QY 2 LPPPRKMKGLFSQAKISLF 20
 DB 176 LPPPRKMKSEAYS-AKIALF 193
 RESULT 27
 AAR91420
 XX AAR91420 standard; protein; 1025 AA.
 XX AAR91420;
 XX 11-JUN-1996 (first entry)
 DT Human dihydropyrimidine dehydrogenase.
 DE Dihydropyrimidine dehydrogenase; DPD; DPD gene; probe; deficiency;
 KW 5-fluorouracil; cytostatic; cancer; antibody.

XX Homo sapiens.
 OS Key Location/Qualifiers
 XX Binding-site /label= NADP(H)_binding_site
 FT Binding-site /label= FAD_binding_site
 FT Binding-site /label= Uracil_binding_site
 FT Binding-site /label= 4Fe-4S_binding_site
 FT Binding-site /label= 4Fe-4S_binding_site
 FT Binding-site /label= 4Fe-4S_binding_site
 XX W09608568-A2.
 XX 21-MAR-1996.
 XX 07-SEP-1995; 95WO-US012016.
 XX 12-SEP-1994; 94US-00304309.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Gonzalez FJ, Fernandez-Salguero P;
 PI WPI; 1996-179943/18.
 DR N-PSDB; AAT14077.
 XX Pig and human dihydro:pyrimidine dehydrogenase (DPD) genes and probes -
 PT useful for detection of DPD deficiencies and identification of humans at
 PT risk of toxic reaction to 5-fluoro:uracil anti-cancer treatment.
 XX Claim 7; Page 48-51; 78pp; English.
 CC Human dihydropyrimidine dehydrogenase (DPD) (AAR91420) can be expressed as
 CC a recombinant protein in a prokaryotic host cell by insertion of an
 CC encoding cDNA sequence (see AAT14077) into a vector, transfection of host
 CC cells, and culturing the cells in medium contg. uracil, FAD and FMN. DPD
 CC is used to raise antibodies useful in competitive binding immunoassays
 XX SQ Sequence 1025 AA;
 Query Match 52.0%; Score 52.5; DB 2; Length 1025;
 Best Local Similarity 63.2%; Pred. No. 15;
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
 QY 2 LPPPRKMKGLFSQAKISLF 20
 DB 176 LPPPRKMKSEAYS-AKIALF 193
 RESULT 28
 AAR93361
 ID AAR93361 standard; protein; 1025 AA.
 XX AAR93361;
 XX 28-MAY-1999 (first entry)
 DT Human DPD protein.
 DE DPD; dihydropyrimidine-dehydrogenase; monoclonal antibody; Mab;
 KW immunoassay reagent; cancer patient; treatment; antitumour agent;
 KW 5-fluorouracil; affinity purification; toxicity.
 XX Homo sapiens.
 XX DE19837391-A1.
 XX 25-FEB-1999.

PF 18-AUG-1998; 98DE-01037391.
PR 22-AUG-1997; 97EP-00114630.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
PI Yoshikubo T, Hasegawa M;
PR WPI; 1999-155202/14.
DR N-PSDB; AAX22903.
XX
PT Monoclonal antibody specific for dihydropyrimidine dehydrogenase - for
PT assessing patient response to 5-fluorouracil antitumor agents.
XX
PS Disclosure; Page 17-22; 34pp; German.
XX
CC This invention describes a monoclonal antibody (MAB) specific for
CC dehydrogenase (DPD). This MAB is used as immunoassay
CC reagents to identify a lack of DPD in a patient and to assess the
CC sensitivity of cancer patients to treatment with antitumor agents of the
CC 5-fluorouracil (5-FU) type. The MAB can also be used for affinity
CC purification of DPD. DPD is involved in reduction of 5-FU (and related
CC catabolites and derivatives) and lack of it is associated with increased
CC toxicity of this type of antitumor agent. It has specific binding
CC interaction. The MAB provide a sensitive and reliable test for DPD, which
CC is simple, rapid and suitable for routine screening
XX
SQ Sequence 1025 AA;

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
Best Local Similarity 63.2%; Pred. No. 15;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPRKMKSEAYS-AKIALF 193
||||| : |||:
||||| : |||:
||||| : |||:

RESULT 29
ADB70177
ID ADB70177 standard; protein; 524 AA.
XX
AC ADB70177;
DT 04-DEC-2003 (first entry)
XX
DE C. neoformans amino acid sequence SEQ ID NO:3221.
XX
KW fungicide; gene therapy; infection.
XX
OS Cryptococcus neoformans.
XX
PN WO2003052076-A2.
XX
PD 26-JUN-2003.
XX
PF 17-DEC-2002; 2002WO-US040225.
XX
PR 17-DEC-2001; 2001US-0341261P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Zamudio C, Eroshkin AM;
XX
DR WPI; 2003-533017/50.
DR N-PSDB; ADB69094.
XX
PT New nucleic acid, useful for preparing a composition for treating an
PT infection caused by Cryptococcus neoformans.
XX
PS Claim 9; SEQ ID NO 3221; 136pp; English.
XX
CC The invention relates to a novel purified or isolated Cryptococcus

CC neoformans nucleic acid molecule comprising a sequence encoding a
CC polypeptide comprising a sequence not given in the specification. A
CC polynucleotide of the invention has nuclease activity, and may have a
CC use in gene therapy. The nucleic acid is useful for preparing a
CC composition for treating an infection caused by Cryptococcus neoformans.
CC The present sequence represents a C. neoformans sequence of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 524 AA;

Query Match 47.5%; Score 48; DB 7; Length 524;
Best Local Similarity 43.8%; Pred. No. 39;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
Qy 2 LPPPRKMKGLFSQAKI 17
Db 5 LPPPRKMKSEAYSQSQW 20
||||| : |||:
||||| : |||:
||||| : |||:

RESULT 30
AAG06085
ID AAG06085 standard; protein; 407 AA.
XX
AC AAG06085;
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2730.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EF1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130492P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135363P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137538P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140822P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142052P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.

PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 08-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161040P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.

PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 23-OCT-1999; 99US-0162142P.

Query Match 46.5%; Score 47; DB 3; Length 407;
Best Local Similarity 56.2%; Pred. NO. 43;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPRKMKGLFSQAKIS 18
|||:|:|:
DB 99 PPRDKQLFKQNVLS 114

Search completed: May 11, 2004, 13:37:32
Job time : 51 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:30:01 ; Search time 10.5 Seconds
(without alignments)
183.222 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR_78:*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	101	100.0	852	1 GNNYHA	genome polyprotein
2	101	100.0	1358	2 A03905	genome polyprotein
3	101	100.0	2227	1 GNNYHM	genome polyprotein
4	101	100.0	2227	1 GNNYHR	genome polyprotein
5	101	100.0	2227	1 GNNYMK	genome polyprotein
6	96	95.0	2230	1 GNNYSA	genome polyprotein
7	95	94.1	2227	1 GNNYHB	genome polyprotein
8	84	83.2	839	1 GNNYS2	genome polyprotein
9	75	74.3	341	2 S04137	dihydropyrimidine
10	52.5	52.0	1025	2 A54718	unknown protein T2
11	47	46.5	414	2 D96838	bacteriophage gene
12	45	44.6	79	2 A01062	probable 60S ribos
13	45	44.6	321	2 E86423	ribonucleoside-dip
14	44	43.6	340	1 WMBES7	hypothetical prote
15	44	43.6	343	2 S74937	C4-dicarboxylase-b
16	44	43.6	369	2 S77299	hypothetical prote
17	44	43.6	775	2 T48957	transposase all756
18	43	42.6	189	2 A12534	methanol dehydroge
19	43	42.6	309	2 D75008	hypothetical prote
20	43	42.6	379	2 T49919	heat shock protein
21	43	42.6	389	2 S41748	hypothetical prote
22	42.5	42.1	105	2 B72598	probable membrane
23	42	41.6	109	2 S69307	hypothetical prote
24	42	41.6	165	2 S02053	ribosomal protein
25	42	41.6	178	2 T02335	hypothetical prote
26	42	41.6	834	2 F82673	hypothetical prote
27	41.5	41.1	304	2 H82984	hypothetical prote
28	41	40.6	83	2 T17809	conserved hypothet
29	41	40.6	159	2 C72210	conserved hypothet

30	41	40.6	213	2 H86638	hypothetical prote
31	41	40.6	229	2 T19153	hypothetical prote
32	41	40.6	237	2 C84065	transcription regu
33	41	40.6	370	2 G37001	endoglucanase fami
34	41	40.6	561	2 G36752	unknown protein F2
35	41	40.6	622	2 T37257	hypothetical prote
36	41	40.6	897	2 T06540	polyribonucleotide
37	41	40.6	1040	2 A34695	axonal glycoprotei
38	41	40.6	1234	2 T30254	jumonji protein -
39	40.5	40.1	924	2 T25007	hypothetical prote
40	40.5	40.1	932	2 T25008	hypothetical prote
41	40	39.6	177	2 C64133	mercuric resistanc
42	40	39.6	238	2 C83240	probable transcrip
43	40	39.6	259	2 F64532	conserved hypothet
44	40	39.6	272	2 T49070	probable heat choc
45	40	39.6	292	2 S23239	hypothetical prote
46	40	39.6	300	2 A81885	hypothetical prote
47	40	39.6	310	2 A55053	endothelial monocy
48	40	39.6	311	2 B55053	endothelial monocy
49	40	39.6	371	2 T46089	thyroid transcript
50	40	39.6	372	2 S53724	thyroid-specific
51	40	39.6	372	2 S12002	thyroid nuclear fa
52	40	39.6	380	2 T20269	hypothetical prote
53	40	39.6	401	2 G02321	thyroid transcript
54	40	39.6	448	2 D64567	glutamate dehydrog
55	40	39.6	523	2 S53945	hypothetical prote
56	40	39.6	562	2 H69545	probable fatty-acid
57	40	39.6	619	2 T03143	R-transactivator p
58	40	39.6	620	2 T50232	actin-like protein
59	40	39.6	645	2 A12963	cellulose synthase
60	40	39.6	645	2 D93319	hypothetical prote
61	40	39.6	649	2 T01106	hypothetical prote
62	40	39.6	739	2 T12964	subtilisin homolog
63	40	39.6	887	2 B96598	hypothetical prote
64	40	39.6	1048	2 T31425	C-terminal domain-
65	39.5	39.1	1040	1 A38306	alpha-mannosidase
66	39	38.6	204	2 D71339	probable ribosomal
67	39	38.6	222	2 G77704	hypothetical prote
68	39	38.6	224	2 A95571	conserved hypothet
69	39	38.6	238	2 E83502	probable transcrip
70	39	38.6	252	2 AB1030	probable membrane
71	39	38.6	274	2 B84847	hypothetical prote
72	39	38.6	280	2 T18443	hypothetical prote
73	39	38.6	347	2 T07108	glucan endo-1,3-be
74	39	38.6	395	2 T32309	hypothetical prote
75	39	38.6	427	2 E72488	probable tryptopa
76	39	38.6	505	2 A95916	probable sugar ABC
77	39	38.6	567	2 C85643	hypothetical prote
78	39	38.6	567	2 H90782	hypothetical prote
79	39	38.6	631	2 A57286	probable serine/th
80	39	38.6	651	2 T31513	hypothetical prote
81	39	38.6	675	2 PQ0227	adenylate cyclase
82	39	38.6	721	2 JC7557	lipidosis - mouse
83	39	38.6	736	2 T12963	subtilisin homolog
84	39	38.6	753	2 T01619	hypothetical prote
85	39	38.6	941	2 AF2415	two-component hybr
86	39	38.6	1127	2 E97580	hypothetical prote
87	39	38.6	1248	2 A53588	adenylate cyclase
88	39	38.6	1251	2 S48687	type VIII adenylyl
89	39	38.6	1258	2 T40737	probable calcium-t
90	39	38.6	1387	2 JCS502	G-protein signalin
91	38.5	38.1	585	2 S55205	dihydroxy-acid deh
92	38.5	38.1	881	2 S67026	probable membrane
93	38	37.6	49	2 F69412	hypothetical prote
94	38	37.6	156	2 B44048	hypothetical 18.1K
95	38	37.6	186	2 T50403	probable succinate
96	38	37.6	190	2 G70231	conserved hypothet
97	38	37.6	192	2 S76867	hypothetical prote
98	38	37.6	223	2 F97120	ribosomal protein
99	38	37.6	248	4 I59064	transforming prote
100	38	37.6	273	2 D69633	glutamine ABC tran

ALIGNMENTS

C:\Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

C:\Accession: A25981
R:\Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.
J. Virol. 61, 50-59, 1987
A:\Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with di
A:\Reference number: A25981; MUID:87061253; PMID:3023706
A:\Accession: A25981
A:\Molecule type: genomic RNA
A:\Residues: 1-2227 <COH>
A:\Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA45465.1; PID:G329583
C:\Superfamily: hepatitis A virus genome polyprotein
C:\Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:\1-23/Product: coat protein 1A #status predicted <VP4>
F:\24-245/Product: coat protein 1B #status predicted <VP2>
F:\246-491/Product: coat protein 1C #status predicted <VP3>
F:\492-836/Product: coat protein 1D #status predicted <VP1>
F:\937-980/Product: coat protein 2A #status predicted <C2A>
F:\981-1087/Product: core protein 2B #status predicted <C2B>
F:\1088-1422/Product: core protein 2C #status predicted <C2C>
F:\1423-1496/Product: core protein 3A #status predicted <C3A>
F:\1497-1519/Product: core protein 3B #status predicted <C3B>
F:\1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:\1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||
DB 823 VLPPPRKMKGLFSQAKISLF 842
|||

RESULT 4

GNNYHR

genome polyprotein - human hepatitis A virus
N:\Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro
NA:\Polymerase (EC 2.7.7.48), protein 3D
C:\Species: human hepatitis A virus
A:\Note: host Homo sapiens (man)
C:\Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
R:\Accession: A03903
R:\Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van Nest
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985
A:\Title: Primary structure and gene organization of human hepatitis A virus.
A:\Reference number: A03903; MUID:85190549; PMID:2986127
A:\Accession: A03903
A:\Molecule type: genomic RNA
A:\Residues: 1-2227 <NAJ>
A:\Cross-references: GB:K02990; NID:G329596; PIDN:AAA45472.1; PID:G329597
C:\Superfamily: hepatitis A virus genome polyprotein
C:\Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans
F:\1-245/Product: coat protein 1A #status predicted <C1A>
F:\246-491/Product: coat protein 1B #status predicted <C1B>
F:\492-836/Product: coat protein 1C #status predicted <C1C>
F:\837-980/Product: core protein 2A #status predicted <C2A>
F:\981-1076/Product: core protein 2B #status predicted <C2B>
F:\1077-1422/Product: core protein 2C #status predicted <C2C>
F:\1423-1484/Product: protein 3A #status predicted <C3A>
F:\1485-1507/Product: protein 3B #status predicted <C3B>
F:\1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>
F:\1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||
DB 823 VLPPPRKMKGLFSQAKISLF 842
|||

RESULT 5

ALIGNMENTS

C:\Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 16-Jul-1999

genome polyprotein - human hepatitis A virus (strain CR326) (fragment)
N:\Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A
C:\Species: human hepatitis A virus
A:\Note: host Homo sapiens (man)
C:\Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Jul-1999
R:\Accession: A03904
R:\Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.M.
J. Virol. 54, 247-255, 1985
A:\Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.
A:\Reference number: A03904; MUID:85185648; PMID:2985793
A:\Accession: A03904
A:\Molecule type: genomic RNA
A:\Residues: 1-852 <LIN>
A:\Cross-references: EMBL:M10033; NID:G329592; PIDN:AAA45470.1; PID:G329593
C:\Superfamily: hepatitis A virus genome polyprotein
C:\Keywords: coat protein; core protein; polyprotein
F:\1-245/Product: coat protein 1A #status predicted <C1A>
F:\246-491/Product: coat protein 1B #status predicted <C1B>
F:\492-836/Product: coat protein 1C #status predicted <C1C>
F:\837-854/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 101; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||
DB 823 VLPPPRKMKGLFSQAKISLF 842
|||

RESULT 2

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)
N:\Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein
C:\Species: human hepatitis A virus
C:\Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 15-Nov-1996
C:\Accession: A03905
R:\Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinston
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985
A:\Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA
A:\Reference number: A03905; MUID:8516289; PMID:2984684
A:\Accession: A03905
A:\Molecule type: genomic RNA
A:\Residues: 1-1358 <BAR>
C:\Superfamily: hepatitis A virus genome polyprotein
C:\Keywords: coat protein; core protein; polyprotein
F:\1-245/Product: coat protein 1A #status predicted <C1A>
F:\246-491/Product: coat protein 1B #status predicted <C1B>
F:\492-836/Product: coat protein 1C #status predicted <C1C>
F:\837-854/Product: core protein 2A (fragment) #status predicted <C2A>
F:\855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 101; DB 2; Length 1358;

Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||
DB 823 VLPPPRKMKGLFSQAKISLF 842
|||

RESULT 3

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)
N:\Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 1D; core pro
B, RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D
C:\Species: human hepatitis A virus
A:\Note: host Homo sapiens (man)

R;Nainan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.
 J. Gen. Virol. 72, 1685-1689, 1991
 A;Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus mac
 A;Reference number: JQ1180; MUID:91311421; PMID:1649902
 A;Accession: JQ1180
 A;Molecule type: genomic RNA
 A;Residues: 1-839 <NAT>
 A;Cross-references: GB:M59286; NID:G329599; PIDN:AAA45473.1; PID:G555083
 C;Superfamily: Hepatitis A virus genome polyprotein
 C;Keywords: coat protein; core protein; glycoprotein; polypeptide
 F;1-23/Product: coat protein 1A #status predicted <VP0>
 F;24-245/Product: coat protein 1B #status predicted <VP3>
 F;246-491/Product: coat protein 1C #status predicted <VP1>
 F;492-839/Product: coat protein 2A (fragment) #status predicted <P2P>
 F;261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.2%; Score 84; DB 1; Length 839;
 Best Local Similarity 83.3%; Pred. No. 8,7e-06;
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKIS 18
 :|||||:|||||:
 Db 822 ILPPPRKMKGLFSQAKIS 839

RESULT 9
 S04137
 genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)
 C;Species: human hepatitis A virus
 C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
 C;Accession: S04137
 R;Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.
 Nucleic Acids Res. 17, 3594, 1989
 A;Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A virus
 A;Reference number: S04137; MUID:89263805; PMID:2542903
 A;Accession: S04137
 A;Molecule type: mRNA
 A;Residues: 1-341 <AND>
 A;Cross-references: EMBL:X14666; NID:G62301; PIDN:CAA32794.1; PID:G4377576
 C;Gene: VP1
 C;Superfamily: hepatitis A virus genome polyprotein
 C;Keywords: coat protein; polypeptide
 F;2-340/Product: coat protein 1D (VP1) #status predicted <NAT>

Query Match 74.3%; Score 75; DB 2; Length 341;
 Best Local Similarity 93.3%; Pred. No. 0.0001;
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQA 15
 :|||||:|||||:
 Db 327 VLPPPRKMKGLFSQS 341

RESULT 10
 A54718
 dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - human
 C;Species: Homo sapiens (man)
 C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 03-Jun-2002
 C;Accession: A54718
 R;Yokota, H.; Fernandez-Salguero, P.; Furuya, H.; Lin, K.; McBride, O.W.; Podschun, B.;
 J. Biol. Chem. 269, 23192-23196, 1994
 A;Title: cDNA cloning and chromosome mapping of human dihydropyrimidine dehydrogenase, a
 A;Reference number: A54718; MUID:94365020; PMID:8083224
 A;Accession: A54718
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1025 <YOK>
 A;Cross-references: GB:U09178; NID:G558304; PID:G558305
 A;Note: authors translated the codon GGC for residue 748 as Asp, and TTA for residue 101
 C;Gene: GDB:DPYD
 A;Cross-references: GDB:364102; OMIM:274270

A;Map position: lp22-1p22
 C;Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] ho
 C;Keywords: 4Fe-4S; flavoprotein; homodimer; iron-sulfur protein; NADP; oxidoreductase
 F;946-1005/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
 Best Local Similarity 63.2%; Pred. No. 2.1;
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGKGLFSQAKISLF 20
 :|||||:|||||:
 Db 176 LPPPRKMSEAYS-AKIALF 193

RESULT 11
 D96838
 unknown protein T21F11.5 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: D96838
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 anken, N.P.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: AB6141; MUID:21016719; PMID:11130712
 A;Accession: D96838
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-414 <STO>
 A;Cross-references: GB:AB005173; NID:G6730725; PIDN:AAF27115.1; GSPDB:GN00141
 C;Gene: T21F11.5
 A;Map position: 1

Query Match 46.5%; Score 47; DB 2; Length 414;
 Best Local Similarity 56.2%; Pred. No. 6.2;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPFRKMKGLFSQAKIS 18
 :|||||:|||||:
 Db 106 PPFRKMKGLFSQAKIS 121

RESULT 12
 AG1062
 bacteriophage gene regulatory protein STY4826 [imported] - Salmonella enterica subsp. en
 C;Species: Salmonella enterica subsp. enterica serovar Typhi
 A;Note: this species has also been called Salmonella typhi
 C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C;Accession: AG1062
 R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Mouton, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A;Reference number: AB0502; MUID:21534947; PMID:11677608
 A;Accession: AG1062
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-79 <PAR>
 A;Cross-references: GB:AL513382; PIDN:CAD06948.1; PID:G16505595; GSPDB:GN00176
 C;Gene: STY4826

Query Match 44.6%; Score 45; DB 2; Length 79;
 Best Local Similarity 63.6%; Pred. No. 2.1;

```

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 PPPRRKMKGLFS 13
Db 64 PPPRRKMGQCS 74

RESULT 13
E86423
probable 60S ribosomal protein L18A - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: E86423
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86423
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-321 <STO>
A:Cross-references: GB:AE005172; NID:G10092460; PIDN:AA12862.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 44.6%; Score 45; DB 2; Length 321;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VLPPIPRKMKGLFSQAKISLF 20
Db 301 VRPPIKLTTPKANKPNLF 320

RESULT 14
WMBE57
ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human herpesvirus 1 (st
N:Alternate names: ribonucleotide reductase small chain
C:Species: human herpesvirus 1
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 15-Sep-2003
C:Accession: D30088
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perz
J. Gen. Virol. 69, 1531-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes sim
A:Reference number: A30083; MUID:88274327; PMID:2839594
A:Accession: D30088
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-340 <MCG>
A:Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32303.1; PID:G59540; GB:D00317
C:Genetics:
A:Gene: UL40
C:Function:
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleos
C:Superfamily: ribonucleoside reductase small subunit
C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; oxid
F:94,124,127,187,221,224/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #sta
F:131/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 43.6%; Score 44; DB 1; Length 340;
Best Local Similarity 61.5%; Pred. No. 16;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPPRRKMKGLFSQA 15
Db 242 PPPRRVYGLFQA 254

```

```

RESULT 15
S74937
hypothetical protein sir0688 - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S74937
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S74937
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-343 <KAN>
A:Cross-references: EMBL:D90902; GB:AB001339; NID:G1652027; PIDN:BAA16977.1; PID:dl01771.1
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 43.6%; Score 44; DB 2; Length 343;
Best Local Similarity 47.1%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPPRRKMKGLFSQAKISL 19
Db 240 PDPRECKGLFRGAFTI 256

RESULT 16
S77299
C4-dicarboxylase-binding protein - Synechocystis sp. (strain PCC 6803)
N:Alternate names: protein sll1314
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S77299
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S77299
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-369 <KAN>
A:Cross-references: EMBL:D90907; GB:AB001339; NID:G1652618; PIDN:BAA17633.1; PID:dl018366
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: dctp

Query Match 43.6%; Score 44; DB 2; Length 369;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 VLPPIPRKMKGLFSQAKISLF 20
Db 38 ILPLPGKVGYYSOAKIRAF 57

RESULT 17
T48957
hypothetical protein T15B3.170 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48957
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Vosse, H.; Unseld, M.; Mewes, H.W.; Rudd, S.; I
submitted to the Protein Sequence Database, April 2000
A:Reference number: 225009
A:Accession: T48957

```

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-775 <JUR>
A;Cross-references: EMBL:AL163975; GSPDB:GN00061; ATSP:T15B3.170
A;Experimental source: cultivar Columbia; BAC clone T15B3
C;Genetics:
A;Gene: ATSP:T15B3.170
A;Map position: 3
A;Introns: 33/3; 135/3; 298/1; 376/2; 478/3; 551/3

Query Match 43.6%; Score 44; DB 2; Length 775;
Best Local Similarity 52.9%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 3 PPRKMKGLFSQAKISL 19
||| ||| ||| |||
Db 749 PPRKMKHLEFNQCHTSL 765

RESULT 18
AI2534
transposase all7564 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120beta
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AI2534
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AI2534
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-189 <KUR>
A;Cross-references: GB:AP003602; PIDN:BAE77207.1; PID:g17134649; GSPDB:GN00181
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all7564
A;Genome: plasmid

Query Match 42.6%; Score 43; DB 2; Length 189;
Best Local Similarity 55.6%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLFSQAKISL 19
||| ||| ||| |||
Db 162 LTPPRKMKNLQWRKILL 179

RESULT 19
D75008
methanol dehydrogenase regulatory protein (moxr-2) PAB1296 - Pyrococcus abyssi (strain C
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C;Accession: D75008
R;anonymous. Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: D75008
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-309 <NAW>
A;Cross-references: GB:A7248288; GB:AL096836; NID:95458960; PIDN:CAB50506.1; PID:g545902
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1296
C;Superfamily: methanol dehydrogenase regulatory protein

Query Match 42.6%; Score 43; DB 2; Length 309;
Best Local Similarity 81.8%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 RKMKGFLFSQAK 16
||| ||| ||| |||
Db 185 RKDKGLFSQAK 195

RESULT 20
T49919
hypothetical protein F17114.20 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49919
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft,
submitted to the Protein Sequence Database, April 2000
A;Reference number: Z24490
A;Accession: T49919
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-379 <BEV>
A;Cross-references: EMBL:AL353994; GSPDB:GN00063; ATSP:F17114.20
A;Experimental source: cultivar Columbia; BAC clone F17114
C;Genetics:
A;Gene: ATSP:F17114.20
A;Map position: 5
A;Introns: 119/1; 146/1; 222/3; 256/2; 327/2

Query Match 42.6%; Score 43; DB 2; Length 379;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 PPRKMKGL 11
||| ||| ||| |||
Db 32 PPRKMKSM 40

RESULT 21
S41748
heat shock protein dnaJ - Methanosarcina mazei
C;Species: Methanosarcina mazei
C;Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 20-Aug-1999
C;Accession: S41748; S41150
R;Conway de Macario, E.
submitted to the EMBL Data Library, April 1992
A;Reference number: S41748
A;Accession: S41748
A;Molecule type: DNA
A;Residues: 1-389 <CON>
A;Cross-references: EMBL:X60265; NID:g48938; PIDN:CAA42813.1; PID:g48940
A;Experimental source: strain S-6
R;Macario, A.J.L.; Dugan, C.B.; Clarens, M.; Conway de Macario, E.
Nucleic Acids Res. 21, 2773, 1993
A;Title: dnaJ in Archaea.
A;Reference number: S41150; MUID:93324351; PMID:8332479
A;Accession: S41150
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 4-208 <MAC>
A;Cross-references: EMBL:X60265; NID:g48938; PIDN:CAA42813.1; PID:g48940
A;Experimental source: strain S-6
C;Genetics:

A;Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology
C;Keywords: DNA replication; heat shock; molecular chaperone; stress-induced protein
F;6-70/Domain: dnaJ amino-terminal homology <DNJ>
F;77-104/Region: G/F motif
F;144-151/Region: CXXCXGKG repeat
F;161-168/Region: CXXCXGKG repeat
F;187-194/Region: CXXCXGKG repeat
F;201-208/Region: CXXCXGKG repeat

Query Match 42.6%; Score 43; DB 2; Length 389;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 5 PRKMKGLFSQAK 16
   ||| ||| :
Db 373 PRKSGLFKPKV 384

RESULT 22
B72598
Hypothetical protein APE1250 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:Accession: B72598
R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382956
A:Accession: B72598
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-105 <KAW>
A:Cross-references: DBJ:AP000061; NID:G5104821; PIDN:BA080240.1; PID:d1044026; PID:G510
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1250

Query Match 42.1%; Score 42.5; DB 2; Length 105;
Best Local Similarity 61.1%; Pred. NO. 7.6;
Matches 11; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 VLPPrPKMKGLFSQAKIS 18
   ||| ||| :
Db 36 VLPPrPKRAM-SMASQGLIS 52

RESULT 23
S69307
probable membrane protein YLR294c - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein I8003.19-a
C:Species: Saccharomyces cerevisiae
C:Date: 20-Jul-1996 #sequence_revision 23-Aug-1996 #text_change 19-Apr-2002
C:Accession: S69307
R:Pauley, A.
submitted to the EMBL Data Library, November 1994
A:Description: The sequence of S. cerevisiae cosmid 8003.
A:Reference number: S50366
A:Accession: S69307
A:Molecule type: DNA
A:Residues: 1-109 <PAU>
A:Cross-references: EMBL:U17243; NID:G596030; PID:G2340967; GSPDB:GN00012; MIPS:YLR294c
C:Genetics:
A:Gene: MIPS:YLR294c
A:Cross-references: SGD:S0004285
A:Map position: 12R
C:Superfamily: Saccharomyces probable membrane protein YLR294c
C:Keywords: transmembrane protein
P:77-93/Domain: transmembrane #status predicted <TM>

Query Match 41.6%; Score 42; DB 2; Length 109;
Best Local Similarity 50.0%; Pred. NO. 9.7;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKISLF 20
   ||| ||| :
Db 63 PHKQNLFFPKQIKLY 78

RESULT 24
S02053
hypothetical protein (D10 5' region) - phage T5
C:Species: phage T5
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 08-Oct-1999
C:Accession: S02053
```

```
R:Kallman, A.V.; Kryukov, V.M.; Bayev, A.A.
Nucleic Acids Res. 16, 10353-10354, 1988
A:Title: The nucleotide sequence of the region of bacteriophage T5 early genes D10-D15.
A:Reference number: S01931; MUID:89057468; PMID:3057441
A:Accession: S02053
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-165 <KAL>
A:Cross-references: EMBL:X12930; NID:G15407; PIDN:CAA31397.1; PID:G579170
C:Genetics:
A:Start codon: GTG
C:Superfamily: phage T5 hypothetical protein (D10 5' region)

Query Match 41.8%; Score 42; DB 2; Length 165;
Best Local Similarity 53.3%; Pred. NO. 15;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 6 RKMKGGLFSQAKISLF 20
   ||| ||| :
Db 38 RKLPGTFSQRRLIF 52

RESULT 25
T02335
ribosomal protein L18a, cytosolic - Arabidopsis thaliana
N:Alternate names: protein F13P17.31
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001
C:Accession: T02335; B84757
R:Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul
submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome II BAC F13P17 genomic sequence.
A:Reference number: Z14657
A:Accession: T02335
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <ROU>
A:Cross-references: EMBL:AC004481; NID:G3337347; PIDN:AAC27421.1; PID:G3337376
A:Experimental source: Cultivar Columbia
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: B84757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-178 <STO>
A:Cross-references: GB:AE002093; NID:G3128228; PIDN:AAC26708.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g34480; F13P17.31
A:Map position: 2
A:Introns: 5/3; 44/2; 65/3
C:Superfamily: rat ribosomal protein L18a
C:Keywords: cytosol; protein biosynthesis; ribosome

Query Match 41.8%; Score 42; DB 2; Length 178;
Best Local Similarity 45.0%; Pred. NO. 17;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 VLPPrPKMKGLFSQAKISLF 20
   ||| ||| :
Db 158 VRPPSRKLTKTYKANKPNLF 177

RESULT 26
T2673
hypothetical protein XF1508 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: T2673
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
```

Nature 406, 151-157, 2000
 A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: F82673
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-834 <STM>
 A/Cross-references: GB:AE003980; GB:AE003849; NID:G9106531; PIDN:AAF84317.1; GSPDB:GN001
 A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; B
 Briones, M.R.S.; Bueno, M.R.F.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigz
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A/Gene: XF1508

Query Match 41.6%; Score 42; DB 2; Length 834;
 Best Local Similarity 60.0%; Pred. No. 93;
 Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 QY 1 VLPPEKMKGLFSQA 15
 DB 597 VAPPPFVKTLRQA 611

RESULT 27
 H82984
 hypothetical protein PA5284 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: H82984
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A>Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: H82984
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-304 <STO>
 A/Cross-references: GB:AE004941; GB:AE004091; NID:G9951596; PIDN:AG08669.1; GSPDB:GN001
 A/Experimental source: strain PAO1
 C/Genetics:
 A/Gene: PA5284

Query Match 41.1%; Score 41.5; DB 2; Length 304;
 Best Local Similarity 50.0%; Pred. No. 37;
 Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;
 QY 2 LPPPRKMKGLFSQ-AKISLF 20
 DB 102 IPASSRFKGTFSQIAEIQLF 121

RESULT 28
 T17809
 hypothetical protein a311R - Chlorella virus PBCV-1
 C/Species: Chlorella virus PBCV-1
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C/Accession: T17809
 R/Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999
 A/Reference number: Z18806
 A/Accession: T17809
 A/Status: preliminary; translated from GB/EMBL/DBDJ
 A/Molecule type: DNA
 A/Residues: 1-83 <GRA>
 A/Cross-references: EMBL:U42580; NID:G4028896; PIDN:AA096679.1
 A/Experimental source: specific host Chlorella strain NC64A
 C/Genetics:
 A/Note: a311R

Query Match 40.6%; Score 41; DB 2; Length 83;
 Best Local Similarity 47.1%; Pred. No. 10;
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 2 LPPPRKMKGLFSQAKIS 18
 DB 50 LPVPAIDKGFASAKLT 66

RESULT 29
 C72210
 conserved hypothetical protein - Thermotoga maritima (strain MSB8)
 C/Species: Thermotoga maritima
 C/Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C/Accession: C72210
 R/Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.
 C.M.
 Nature 399, 323-329, 1999
 A>Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A/Reference number: A72200; MUID:99287316; PMID:10360571
 A/Accession: C72210
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-159 <ARN>
 A/Cross-references: GB:AE001817; GB:AE000512; NID:G4982370; PIDN:AA036861.1; PID:G498237
 A/Experimental source: strain MSB8
 C/Genetics:
 A/Gene: TM1798
 C/Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0377

Query Match 40.6%; Score 41; DB 2; Length 159;
 Best Local Similarity 44.4%; Pred. No. 22;
 Matches 8; Conservative 2; Mismatches 8; Indels 0; Gaps 0;
 QY 3 PPPPRKMKGLFSQAKISLF 20
 DB 139 PPPPRKMKGLFSQAKISLF 156

RESULT 30
 H86638
 hypothetical protein ybbe [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
 C/Species: Lactococcus lactis subsp. lactis
 C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
 C/Accession: H86638
 R/Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrli
 Genome Res. 11, 731-753, 2001
 A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s
 A/Reference number: A86625; MUID:21235186; PMID:11337471
 A/Accession: H86638
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-213 <STO>
 A/Cross-references: GB:AE005176; PID:G12722959; PIDN:AA04210.1; GSPDB:GN00146
 A/Experimental source: strain IL1403
 C/Genetics:
 A/Gene: ybbe

Query Match 40.6%; Score 41; DB 2; Length 213;
 Best Local Similarity 46.7%; Pred. No. 30;
 Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 402 IPASSRFKGTFSQIAEIQLF 121

Qy 5 PRKMKGLFSQAKISL 19
Db | : : : : :
153 PLTMKGVPKQKYSI 167

Search completed: May 11, 2004, 13:41:21
Job time : 13.75 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:24:26 ; Search time 6.75 Seconds
(without alignments)
154.282 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101
Sequence: 1 VLPFRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	852	1	POLG HPACV
2	101	100.0	2227	1	POLG HPACV
3	101	100.0	2227	1	POLG HPACV
4	96	95.0	2226	1	POLG HPACV
5	96	95.0	2226	1	POLG HPACV
6	96	95.0	2226	1	POLG HPACV
7	95	95.0	2230	1	POLG HPACV
8	95	94.1	2227	1	POLG HPACV
9	84	83.2	839	1	POLG HPACV
10	75	74.3	341	1	POLG HPACV
11	52.5	52.0	1025	1	DPYD BOVIN
12	45.5	45.0	1025	1	DPYD BOVIN
13	45	44.6	178	1	RLIX ORYSA
14	45	44.6	178	1	RLIX ARATH
15	44	43.6	340	1	RIR2 HSV11
16	44	43.6	340	1	RIR2 HSV1K
17	43	42.6	389	1	DNAJ METMA
18	43	42.6	597	1	MTSL STRAL
19	42	41.6	178	1	RLIX ARATH
20	41	40.6	289	1	NK26 MOUSE
21	41	40.6	359	1	MCAL CRIGR
22	41	40.6	1040	1	AXO1 RAT
23	41	40.6	1234	1	JMJ MOUSE
24	40	39.6	135	1	ZNF7 HAEIN
25	40	39.6	292	1	YOM1 CAEL
26	40	39.6	310	1	MCAL MOUSE
27	40	39.6	312	1	MCAL HUMAN
28	40	39.6	353	1	NK24 HUMAN
29	40	39.6	354	1	NK24 MOUSE
30	40	39.6	371	1	TTFL CANPA
31	40	39.6	371	1	TTFL HUMAN
32	40	39.6	372	1	TTFL MOUSE
33	40	39.6	372	1	TTFL RAT

34	40	39.6	448	1	DHE4 YEAST
35	40	39.6	523	1	YMR9 YEAST
36	40	39.6	620	1	ARP8 SCHPO
37	40	39.6	659	1	SYM ENTPA
38	40	39.6	1048	1	SRA1 RAT
39	40	39.6	1157	1	SRA4 HUMAN
40	39.5	39.1	1039	1	M2C1 MOUSE
41	39.5	39.1	1040	1	M2C1 RAT
42	39	38.6	178	1	RL1X CASSA
43	39	38.6	204	1	RS4 TRSPA
44	39	38.6	222	1	Y039 RICCN
45	39	38.6	347	1	EL1A SOYBN
46	39	38.6	427	1	TRB1 ARPE
47	39	38.6	462	1	WDR8 MOUSE
48	39	38.6	615	1	CNK RAT
49	39	38.6	631	1	CNK MOUSE
50	39	38.6	646	1	CNK HUMAN
51	39	38.6	673	1	SYM OCEIH
52	39	38.6	808	1	POLG HPAVG
53	39	38.6	957	1	IF2 SYNEL
54	39	38.6	1103	1	DPOD RAT
55	39	38.6	1107	1	KCH8 HUMAN
56	39	38.6	1248	1	CYAS RAT
57	39	38.6	1249	1	CYAS MOUSE
58	39	38.6	1251	1	CYAS HUMAN
59	39	38.6	1387	1	RGSC RAT
60	38.5	38.1	585	1	ILV3 YEAST
61	38.5	38.1	881	1	ALPS YEAST
62	38	37.6	49	1	YD03 ARCFU
63	38	37.6	203	1	LOLE VIBPA
64	38	37.6	233	1	RS2 CLOAB
65	38	37.6	317	1	WR15 ARATH
66	38	37.6	398	1	MPK2 CHICK
67	38	37.6	429	1	FXP3 MOUSE
68	38	37.6	465	1	HN4A HUMAN
69	38	37.6	465	1	HN4A MOUSE
70	38	37.6	491	1	HN4A RAT
71	38	37.6	491	1	YA29 SCHPO
72	38	37.6	496	1	MGLA TREPA
73	38	37.6	516	1	ROBG MOUSE
74	38	37.6	752	1	CLPE STRPN
75	38	37.6	1019	1	DLF2 HUMAN
76	38	37.6	1089	1	DLF2 RAT
77	38	37.6	1223	1	K125 CAEL
78	38	37.6	1266	1	JMJ HUMAN
79	38	37.6	1434	1	VG65 HSV11
80	38	37.6	1787	1	CHD3 CAEL
81	38	37.6	2839	1	NFI HUMAN
82	37.5	37.1	329	1	GC3 MOUSE
83	37.5	37.1	398	1	GC3M MOUSE
84	37.5	37.1	430	1	CINA MYCTU
85	37	36.6	173	1	YN26 CAEL
86	37	36.6	261	1	RL1 HSV2H
87	37	36.6	293	1	US17 HCMVA
88	37	36.6	301	1	ORT1 HUMAN
89	37	36.6	312	1	O2L2 HUMAN
90	37	36.6	315	1	LIP3 MORSP
91	37	36.6	356	1	MPCP RAT
92	37	36.6	357	1	MPCP MOUSE
93	37	36.6	364	1	MRP AQUAE
94	37	36.6	379	1	TSY1 MOUSE
95	37	36.6	392	1	OR9A DROME
96	37	36.6	413	1	FL1 TOBAC
97	37	36.6	428	1	ACOF XENNE
98	37	36.6	450	1	SR5A EUCBP
99	37	36.6	468	1	PER1 RAT
100	37	36.6	469	1	SYE1 THEMA

ALIGNMENTS

RESULT 1

POLG HPAVC STANDARD; PRT; 852 AA.

AC P06442; Q83741; Q83742;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).
 DE Hepatitis A virus (strain CR326).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatocovirus.
 OX NCBI_TaxID=12097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85185648; PubMed=2985793;
 RA Lineneyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,
 RA Young A., Mitra S.W.;
 RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA";
 RL J. Virol. 54:247-255(1985).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M10033; AAA45470.1; -;
 CC PIR; A03904; GNNYHA.
 CC InterPro; IPR008975; Viral_cap_coat.
 KW Polyprotein; Coat protein; Core protein.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
 FT CHAIN 837 >852 CORE PROTEIN P2A.
 FT NON TER 852 852
 SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820B CRC64;
 Query Match 100.0%; Score 101; DB 1; Length 852;
 Best Local Similarity 100.0%; Pred. No. 9.9e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 823 VLPPPRKMKGLFSQAKISLF 842
 RESULT 2
 ID POLG HPAVC STANDARD; PRT; 2227 AA.
 AC P06617; P06443; Q81082;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
 DE Hepatitis A virus (strain HM-175).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatocovirus.
 OX NCBI_TaxID=12098;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Wild type;
 RX MEDLINE=87061253; PubMed=3023706;
 RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,

Baroudy B.M.;
 "Complete nucleotide sequence of wild-type hepatitis A virus:
 comparison with different strains of hepatitis A virus and other
 picornaviruses";
 J. Virol. 61:50-59(1987).
 [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Attenuated;
 RX MEDLINE=87175701; PubMed=3031886;
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M.,
 RA Purcell R.H.;
 RT "Complete nucleotide sequence of an attenuated hepatitis A virus:
 comparison with wild-type virus";
 Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
 [3]
 RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
 RX MEDLINE=85166289; PubMed=2984684;
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Maizel J.V. Jr.,
 RA Purcell R.H., Feinstein S.M.;
 RT "Sequence analysis of hepatitis A virus cDNA coding for capsid
 RT proteins and RNA polymerase";
 Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA) (N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
 CC SHOWN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M14114; AAA45475.1; -;
 CC EMBL; M14707; AAA45465.1; -;
 CC EMBL; M14707; AAA45466.1; ALT_INIT.
 CC PIR; A03905; A03905.
 CC PIR; A25981; GNNYHM.
 CC PIR; A94149; GNNYMK.
 CC PDB; 1HAV; 23-DEC-96.
 CC MEROPS; C03.005; -;
 CC InterPro; IPR004004; Calici_pol_hel.
 CC InterPro; IPR009003; Cys_Ser_trypsin.
 CC InterPro; IPR006005; RNA_helicase.
 CC InterPro; IPR007095; RNA_pol_DS_PS.
 CC InterPro; IPR001205; RNA_pol_P3D.
 CC InterPro; IPR007094; RNA_pol_P5vir.
 CC InterPro; IPR008975; Viral_cap_coat.
 CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam; PF00910; RNA_helicase; 1.
 CC PRINTS; PR00918; CALICIVIRUSN.
 KW Polyprotein; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23
 FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).
 FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).
 FT CHAIN 492 836 COAT PROTEIN VP3 (P1C).
 FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).
 FT CHAIN 981 1087 CORE PROTEIN P2A.
 FT CHAIN 1088 1422 CORE PROTEIN P2B.
 FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.


```

DR PRINTS; PR00918; CALICIVIRUSNS.
DR Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
DR CHAIN 1 23
DR FT CHAIN 24 245
DR FT CHAIN 246 491
DR FT CHAIN 432 836
DR FT CHAIN 837 980
DR FT CHAIN 981 1076
DR FT CHAIN 1077 1422
DR FT CHAIN 1423 1484
DR FT CHAIN 1485 1507
DR FT CHAIN 1508 1678
DR FT CHAIN 1679 2227
DR SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
||| ||||| ||||| ||||| |||||
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 4
POLG_HPAV2 STANDARD; PRT; 2226 AA.
ID POLG_HPAV2 AC P26580;
DC 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12094;
RX [1]
SEQUENCE FROM N.A.
MEDLINE=91162758; PubMed=1705995;
Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
Cromeans T., Jansen R.W.;
"Antigenic and genetic variation in cytopathic hepatitis A virus
variants arising during persistent infection: evidence for genetic
recombination.";
J. Virol. 65:2056-2065(1991).
CC -!- CATALYTIC ACTIVITY; N nucleoside triphosphate = N diphosphate +
(RNA)(N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).
CC EMBL; M59810; AAA45468.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR006005; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSwir.
DR InterPro; IPR008975; Viral_cap_coat.

```

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B0BF75 CRC64;
Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842
RESULT 6
POLG_HPAV8 STANDARD; PRT; 2226 AA.
ID AC P26582;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48).
OS Hepatitis A virus (strain 18f).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12096;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; M59809; AAA45467.1; --
DR MEROPS; C03.005; --
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR009003; Cys_Ser_trypsin.

DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;
Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842
RESULT 5
POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
DE P3D (EC 2.7.7.48).
OS Hepatitis A virus (strain 43c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.";
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA}(N).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL; M59809; AAA45469.1; --
DR MEROPS; C03.005; --
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR009003; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D00924; BAA00766.1; -;
 EMBL; X15461; CAA33490.1; -;
 PIR; A30470; GNNYSA.
 MEROPS; C03.005; -;
 InterPro; IPR004004; Calici_pol_hel.
 InterPro; IPR009003; Cys_Ser_trypsin.
 InterPro; IPR006005; RNA_helicase.
 InterPro; IPR007095; RNA_pol_DS_PS.
 InterPro; IPR007095; RNA_pol_P3D.
 InterPro; IPR001205; RNA_pol_P3D.
 InterPro; IPR007094; RNA_pol_PSVir.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00680; RNA_dep_RNA_pol; 1.
 Pfam; PF00910; RNA_helicase; 1.
 PRINTS; PR00918; CALICIVIRUSNS.
 Polyprotein; Coat protein; Core protein; Transferase;
 RNA-directed RNA polymerase; Hydrolase; Thiol protease.
 FT CHAIN 1 27 COAT PROTEIN VP4 (P1A).
 FT CHAIN 28 249 COAT PROTEIN VP2 (P1B).
 FT CHAIN 250 495 COAT PROTEIN VP3 (P1C).
 FT CHAIN 496 795 COAT PROTEIN VP1 (P1D).
 FT CHAIN 796 984 CORE PROTEIN P2A.
 FT CHAIN 985 1091 CORE PROTEIN P2B.
 FT CHAIN 1092 1426 CORE PROTEIN P2C.
 FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
 FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
 FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
 FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
 SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F9 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2230;
 Best Local Similarity 90.0%; Pred. No. 1.8e-07;
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
 :||||:||||:||||:||||:
 Db 827 ILPPPRKVKGLFSQAKISLF 846

RESULT 8
 POLG HPAAV STANDARD; PRT; 2227 AA.
 AC P1390; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;
 AC Q81090; Q81091; Q81092; Q81093;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Hepatitis A virus (strain MB).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12100;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=88045071; PubMed=2823500;
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E.,
 RA Deinhardt F.;
 RA "The entire nucleotide sequence of the genome of human hepatitis A
 RT virus (isolate MB).";
 RL Virus Res. 8:153-171(1987).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT

InterPro; IPR000605; RNA_helicase.
 InterPro; IPR007095; RNA_pol_DS_PS.
 InterPro; IPR001205; RNA_pol_P3D.
 InterPro; IPR007094; RNA_pol_PSVir.
 InterPro; IPR008975; Viral_cap_coat.
 Pfam; PF00680; RNA_dep_RNA_pol; 1.
 Pfam; PF00910; RNA_helicase; 1.
 PRINTS; PR00918; CALICIVIRUSNS.
 Polyprotein; Coat protein; Core protein; Transferase;
 RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
 FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
 FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
 FT CHAIN 795 900 CORE PROTEIN P2A.
 FT CHAIN 901 1087 CORE PROTEIN P2B.
 FT CHAIN 1088 1422 CORE PROTEIN P2C.
 FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.
 FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
 FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
 FT CHAIN 1738 2226
 SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
 Best Local Similarity 95.0%; Pred. No. 1.8e-07;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
 :||||:||||:||||:||||:
 Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 7
 POLG HPAAV STANDARD; PRT; 2230 AA.
 AC P14553;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
 DE P3D (EC 2.7.7.48)].
 OS Simian hepatitis A virus (strain AGM-27).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OX NCBI_TaxID=12102;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91311420; PubMed=1649901;
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,
 RA Purcell R.H.;
 RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
 RT structure and growth in cell culture with other HAV strains.";
 RL J. Gen. Virol. 72:1677-1683(1991).
 RN [2]
 RP SEQUENCE OF 1750-2164 FROM N.A.
 RA MEDLINE=89232168; PubMed=2541023;
 RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
 RA Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
 RA "Variations in genome fragments coding for RNA polymerase in human
 RT and simian hepatitis A viruses.";
 RL FEBS Lett. 247:425-428(1989).
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 CC (RNA)(N).
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
 CC each of which is composed of one copy each of proteins VP1, VP2,
 CC VP3, and VP4.
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20273; AAA45474.1; -
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_ser_trypsin.
CC InterPro; IPR009605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_P5vir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSN.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 837 980 COAT PROTEIN VP1 (PID).
FT CHAIN 981 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2227 AA; 251435 MW; EC983ED2A7C86349 CRC64;

Query Match 94.1%; Score 95; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 9
POLG HPVAVT STANDARD; PRT; 839 AA.
AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20273; AAA45474.1; -
CC MEROPS; C03.005; -
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_ser_trypsin.
CC InterPro; IPR009605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_P5vir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSN.
CC PolyProtein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 837 980 COAT PROTEIN VP1 (PID).
FT CHAIN 981 1087 CORE PROTEIN P2A.
FT CHAIN 1088 1422 CORE PROTEIN P2B.
FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
SQ SEQUENCE 2227 AA; 251435 MW; EC983ED2A7C86349 CRC64;

Query Match 94.1%; Score 95; DB 1; Length 2227;
Best Local Similarity 95.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 9
POLG HPVAVT STANDARD; PRT; 839 AA.
AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
DE P2A] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M59286; AAA45473.1; -
CC PIR; JQ1180; GNNYS2.
CC InterPro; IPR008975; Viral_cap_coat.
CC PolyProtein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 6.2e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKIS 18
DB 822 ILPPPRKMKGLFSQSKIS 839

RESULT 10
POLG HPVAVT STANDARD; PRT; 341 AA.
AC P13672;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP3 (1C); Coat protein VP1
DE (1D); Core protein P2A] (Fragment).
OS Hepatitis A virus (strain LCDG-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14666; CAA32794.1; -
CC PIR; S04137; S04137.
CC InterPro; IPR008975; Viral_cap_coat.
CC PolyProtein; Coat protein; Core protein.
FT NON_TER 1 1 COAT PROTEIN VP3.
FT CHAIN <1 1 COAT PROTEIN VP1.
FT CHAIN 2 340 COAT PROTEIN P2A.
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF136D5 CRC64;

Query Match 74.3%; Score 75; DB 1; Length 341;


```

FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT VARIANT 29 29 C -> R (in allele DPYD*9A and allele
FT DPYD*9B; loss of activity;
FT dbSNP:1801265).
FT /FTID=VAR_005173.
FT R -> W (in allele DPYD*8; loss of
FT activity; dbSNP:1801266).
FT /FTID=VAR_005174.
FT S -> N (in allele DPYD*4; low activity;
FT dbSNP:1801158).
FT /FTID=VAR_005175.
FT I -> V (in allele DPYD*5; dbSNP:1801159).
FT /FTID=VAR_005176.
FT V -> I (in dbSNP:1801160).
FT /FTID=VAR_014760.
FT R -> H (in allele DPYD*9B; 25% of
FT activity; dbSNP:1801267).
FT /FTID=VAR_005177.
FT V -> F (in allele DPYD*10; low activity;
FT dbSNP:1801268).
FT /FTID=VAR_005178.
FT S -> N (IN REF. 2 AND 3).
FT /FTID=VAR_005179.
FT CONFLICT 910 910 521C9430C7F69AFA CRC64;
FT SEQUENCE 1025 AA; 111374 MW; 521C9430C7F69AFA CRC64;
FT
FT Query Match 52.0%; Score 52.5; DB 1; Length 1025;
FT Best Local Similarity 63.2%; Pred. No. 1.2;
FT Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
FT
QY 2 LPPPRKMGKGLFSQAKISL 20
DB 176 LPPPEKMGSEAYS-AKIALF 193
||||| : |||||
[2]

RESULT 12
DPYD BOVIN
ID DPYD BOVIN STANDARD; PRT; 1025 AA.
AC Q28007; QSTRV4;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Dihydropyrimidine dehydrogenase [NADP+] (EC 1.3.1.2) (DPD)
DE (DHPDHase) (Dihydrouracil dehydrogenase) (Dihydrothymine
DE dehydrogenase).
DE
GN DPYD.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
OX [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=97069948; PubMed=8912928;
RX Albin N., Johnson M.R., Diasio R.B.;
RA "cDNA cloning of bovine liver dihydropyrimidine dehydrogenase.";
RT DNA Seq. 6:243-250(1996).
RN [2]
RP SEQUENCE OF 668-678, AND ENZYME REGULATION.
RC TISSUE=Liver;
RX MEDLINE=92041818; PubMed=1939061;
RA Porter D.J.T., Chestnut W.G., Taylor L.C.E., Merrill B.M., Spector T.;
RT "Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.";
RL J. Biol. Chem. 266:19988-19994(1991).
CC -|- FUNCTION: Involved in pyrimidine base degradation. Catalyzes the
CC reduction of uracil and thymine.
CC -|- CATALYTIC ACTIVITY: 5,6-dihydrouracil + NADP(+) = uracil + NADPH.
CC -|- COFACTOR: FAD and FMN. Also contains 2 4Fe-4S clusters.
CC -|- ENZYME REGULATION: Inactivated by 5-iodouracil.
CC -|- PATHWAY: Initial and rate-limiting enzyme in the 3-step pathway of
uracil and thymidine catabolism and in the pathway leading to the
formation of beta-alanine.
-|- SUBUNIT: Homodimer (By similarity).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/ebis-sib.ch/
or send an email to license@ebi-sib.ch).
-----
CC EMBL; U20981; AAB40985.1; -
DR InterPro; IPR001450; 4Fe4S_ferredoxin.
DR InterPro; IPR000759; Adrnx_reductase.
DR InterPro; IPR001295; DHO dh.
DR InterPro; IPR005720; DHO dh1.
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR003009; FMN_enzyme.
DR InterPro; IPR001100; Pyr_redox.
DR InterPro; IPR00103; Pyridine_redox_2.
DR Pfam; PF01180; DHODEHase; 1.
DR Pfam; PF00037; fer4; 2.
DR Pfam; PF00070; Pyr_redox; 1.
DR PRINTS; PR00419; ADXRDXTASE.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; ENDRDTASEII.
DR PRINTS; PR00469; ENDRDTASEII.
DR TIGRfams; TIGR01037; pyrd sub1 fam; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
KW Oxidoreductase; NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4Fe-4S.
FT METAL 953 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT SEQUENCE 1025 AA; 111696 MW; 5B55F93A06C47F4F CRC64;
FT
FT Query Match 45.0%; Score 45.5; DB 1; Length 1025;
FT Best Local Similarity 61.1%; Pred. No. 17;
FT Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
FT
QY 2 LPPPRKMGKGLFSQAKISL 19
DB 176 LPPPEKMGSEAYS-AKIAL 192
||||| : |||||
[2]

RESULT 13
RLIX ORYSA
ID RLIX ORYSA STANDARD; PRT; 178 AA.
AC Q943F3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 60S ribosomal protein L18a.
GN RPL18A.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
RT clone: P0046805.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AP003237; BAB67920.1; --
DR EMBL; AP003249; BAB89536.1; --
DR Gramene; Q943F3; --
DR InterPro; IPR002670; Ribosomal_L18ae.
DR Pfam; PF01775; Ribosomal_L18ae; 1.
KW Ribosomal protein.
SQ SEQUENCE 178 AA; 21400 MW; 1A508AA670C8D93B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;
Best Local Similarity 45.0%; Pred. No. 3.2;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLPPTKMKGLFSQAKISLF 20
DB 158 VRPTKRLTKTKASRPNLF 177

RESULT 14

ID RL1Y ARATH STANDARD; PRT; 178 AA.
AC Q9LUD4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 60S ribosomal protein L18a-2.
GN RPL18AB OR AT3G14600 OR M181.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Sato S., Nakamura Y., Kareko T., Katoh T., Asamizu E., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty p1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J., Cheuk R.F.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.P.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallander E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
CC -!- SIMILARITY: Belongs to the L18AE family of ribosomal proteins.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; AB023038; BAB02392.1; --
DR EMBL; AV072540; BAE60048.1; --
DR InterPro; IPR002670; Ribosomal_L18ae.
DR Pfam; PF01775; Ribosomal_L18ae; 1.
KW Ribosomal protein.
SQ SEQUENCE 178 AA; 21309 MW; 28A1B900F3C4402B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;
Best Local Similarity 45.0%; Pred. No. 3.2;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLPPTKMKGLFSQAKISLF 20
DB 158 VRPTKRLTKTKASRPNLF 177

RESULT 15

ID RIR2 HSV11 STANDARD; PRT; 340 AA.
AC P10224;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase) (38 kDa subunit).
GN UL40.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.";
RL J. Gen. Virol. 69:1531-1574(1988).
CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioredoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioredoxin.
CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC -!- PATHWAY: DNA replication pathway; first step.
CC -!- SUBUNIT: Heterodimer of a large and a small chain.
CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC small chain family.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; X14112; CAA32303.1; --
DR PIR; D30088; WMBES7.
DR InterPro; IPR000358; Ribonucleotidase.
DR Pfam; PF00268; ribonuc red sm; 1.
DR PROSITE; PS00368; RIBOED SMALL; 1.
KW Oxidoreductase; DNA replication; Metal-binding; Iron.
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 187 187 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT ACT SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 340 AA; 484ED994BF74FD3F CRC64;

Db 242 PPPDRVYGLFRQA 254

RESULT 17

DNAJ METWA

ID DNAJ METWA STANDARD; PRT; 389 AA.

AC P35515;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Chaperone protein dnaJ.

GN DNAJ OR MM2504.

OS Methanosarcina mazei (Methanosarcina frisia).

OC Archaea; Euryarchaeota; Methanomicrobia; Methanosarcinales;

OC Methanosarcinaceae; Methanosarcina.

OX NCBI_TaxID=2209;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S-6;

RX MEDLINE=93324351; PubMed=8332479;

RA Macario A.J.L., Dugan C.B., Clarens M., Conway de Macario E.;

RT "dnaJ in Archaea.";

RL Nucleic Acids Res. 21:2773-2773 (1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;

RX MEDLINE=22120827; PubMed=12125824;

RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,

RA Martinek-Arias R., Henne A., Wietzer A., Baeumer S., Jacobi C.,

RA Brueggenmann H., Lienard T., Christmann A., Soennecke M., Steckel S.,

RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,

RA Fritz H.-J., Gottschalk G.;

RT "The genome of Methanosarcina mazei: evidence for lateral gene transfer between Bacteria and Archaea.";

RL J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).

CC -!- FUNCTION: Acts as a co-chaperone. Stimulates, jointly with grpE, the ATPase activity of dnaK (By similarity).

CC -!- COFACTOR: Binds 2 zinc ions per monomer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: Belongs to the dnaJ family.

CC -!- SIMILARITY: Contains 1 J domain.

CC -!- SIMILARITY: Contains 1 CR domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X60265; CAA42813.1; -.

DR EMBL; AE013494; AAM32200.1; -.

DR PIR; S41748; S41748.

DR HSP; P25685; IHDOJ.

DR InterPro; IPR002939; DnaJ_C.

DR InterPro; IPR001305; DnaJ_CXXCXGXG.

DR InterPro; IPR001623; DnaJ_N.

DR InterPro; IPR008971; HSP40_DnaJ_pep.

DR InterPro; IPR003095; Hsp_DnaJ.

DR Pfam; PF00226; DnaJ; 1.

DR Pfam; PF01556; DnaJ_C; 1.

DR Pfam; PF00684; DnaJ_CXXCXGXG; 1.

DR PRINTS; PR00625; DnaJPROTEIN.

DR SMART; SM00271; DnaJ; 1.

DR PROSITE; PS00636; DnaJ_1; 1.

DR PROSITE; PS00076; DnaJ_2; 1.

DR PROSITE; PS00637; DnaJ_CXXCXGXG; 1.

DR Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding; Complete proteome.

FT DOMAIN 4 72 J-DOMAIN.

FT DOMAIN 73 110 GLY-RICH.

FT REPEAT 144 151 CXXCXGXG MOTIF.

Query Match 43.6%; Score 44; DB 1; Length 340;

Best Local Similarity 61.5%; Pred. No. 9.3;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPPEKMKGLFSQA 15

Db 242 PPPDRVYGLFRQA 254

RESULT 16

RIR2 HSVIK

ID RIR2 HSVIK STANDARD; PRT; 340 AA.

AC P06474;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)

DE (Ribonucleoside reductase) (38 kDa subunit).

GN UL40.

OS Herpes simplex virus (type 1 / strain KOS).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Alphaherpesvirinae; Simplexvirus.

OX NCBI_TaxID=10306;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83059830; PubMed=6292456;

RA Draper K.G., Frink R.J., Wagner E.K.;

RT "Detailed characterization of an apparently unspliced beta herpes simplex virus type 1 gene mapping in the interior of another.";

RL J. Virol. 43:1123-1128 (1982).

RN [2]

RP REVISIONS.

RA Wagner E.K.;

RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: Provides the precursors necessary for DNA synthesis.

CC -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized thioredoxin + H(2)O = ribonucleoside diphosphate + reduced thioredoxin.

CC -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).

CC -!- PATHWAY: DNA replication pathway; first step.

CC -!- SUBUNIT: Heterodimer of a large and a small chain.

CC -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase small chain family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; J02212; AAA66436.1; -.

DR InterPro; IPR000358; Ribonucleotid. redctse.

DR Pfam; PF00268; ribonuc_red_sm_1.

DR PROSITE; PS00368; RIBRED_SNAUL; 1.

DR Oxidoreductase; DNA replication; Metal-binding; Iron.

FT METAL 94 94 IRON 1 (BY SIMILARITY).

FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).

FT METAL 127 127 IRON 1 (BY SIMILARITY).

FT METAL 131 131 BY SIMILARITY.

FT ACT_SITE 187 187 IRON 2 (BY SIMILARITY).

FT METAL 221 221 IRON 2 (BY SIMILARITY).

FT METAL 224 224 IRON 2 (BY SIMILARITY).

SEQUENCE 340 AA; 37966 MW; 921DC04B9D278DE5 CRC64;

Query Match 43.6%; Score 44; DB 1; Length 340;

Best Local Similarity 61.5%; Pred. No. 9.3;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 PPPEKMKGLFSQA 15

FT REPEAT 161 168 CXXCXGXG MOTIF.
 FT REPEAT 187 194 CXXCXGXG MOTIF.
 FT REPEAT 201 208 CXXCXGXG MOTIF.
 FT METAL 144 144 ZINC 1 (BY SIMILARITY).
 FT METAL 147 147 ZINC 2 (BY SIMILARITY).
 FT METAL 161 161 ZINC 2 (BY SIMILARITY).
 FT METAL 164 164 ZINC 2 (BY SIMILARITY).
 FT METAL 187 187 ZINC 2 (BY SIMILARITY).
 FT METAL 190 190 ZINC 2 (BY SIMILARITY).
 FT METAL 201 201 ZINC 1 (BY SIMILARITY).
 FT METAL 204 204 ZINC 1 (BY SIMILARITY).
 SQ SEQUENCE 389 AA; 42990 MW; 5B840E379D8C9139 CRC64;
 Query Match 42.6%; Score 43; DB 1; Length 389;
 Best Local Similarity 66.7%; Pred. No. 16;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 5 PRKMGLEFSQAK 16
 DB 373 PRKMGLEFSQAK 384
 ID MTS1_STRAL STANDARD; PRT; 587 AA.
 AC Q53609;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Modification methylase Sali (EC 2.1.1.72) (Adenine-specific
 DE methyltransferase Sali) (M.Sali).
 GN SALIM.
 OS Streptomyces albus G.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95129852; PubMed=7828868;
 RA Rodicio M.R., Quinton-Jager T., Moran L.S., Slatko B.E., Wilson G.G.;
 RT "Organization and sequence of the Sali restriction-modification
 RT system.";
 RL Gene 151:167-172(1994).
 CC -1- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
 CC GTGAC, CAUSES SPECIFIC METHYLATION ON A-5 ON BOTH STRANDS, AND
 CC PROTECTS THE DNA FROM CLEAVAGE BY THE SALI ENDONUCLEASE.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA adenine = S-
 CC adenosyl-L-homocysteine + DNA 6-methyladenine.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U01232; AAA81887.1; -
 CC HSP; P14385; 2ADM.
 CC REBASE; 3491; M.Sali.
 CC InterPro; IPR02236; N12N6_mcfase.
 CC InterPro; IPR02052; N6_Mtase.
 CC InterPro; IPR000051; SAM_bind.
 CC PRINTS; PR00507; N12N6MTFRASE.
 CC PROSITE; PS00092; N6_MTASE; 1.
 KW Transferase; Methyltransferase; Restriction system.
 SQ SEQUENCE 587 AA; 64975 MW; 5482F374397FF999 CRC64;
 Query Match 42.6%; Score 43; DB 1; Length 587;
 Best Local Similarity 70.0%; Pred. No. 24;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 VLPFPKMGK 10

Db 334 ILPSPKMGK 343
 RESULT 19
 RL1X_ARATH
 ID RL1X_ARATH STANDARD; PRT; 178 AA.
 AC P51418; O64699;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 60S ribosomal protein L18a-1.
 GN RPL18AA OR AT2G34480 OR T31810.18.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H.L.,
 RA Moffat K.S., Cronin L.A., Shen M., Pal G., Van Aken S., Unayam L.,
 RA Tallon L.J., Gill J.E., Adams M.D., Carrera A.J., Creasy T.H.,
 RA Goodman H.M., Somerville C.R., Copenhaver G.P., Preuss D.,
 RA Niernan W.C., White O., Bisen J.A., Salzberg S.L., Fraser C.M.,
 RA Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=22954850; PubMed=14593172;
 RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
 RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
 RA Karlyshin N.M., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
 RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
 RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
 RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
 RA Khan S., Koseena E., Ishida J., Jiang P.X., Jones T., Kawai J.,
 RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
 RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
 RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
 RT "Empirical analysis of transcriptional activity in the Arabidopsis
 RT genome.";
 RN Science 302:842-846(2003).
 RN [3]
 RP SEQUENCE OF 69-160 FROM N.A.
 RC STRAIN=cv. Columbia;
 RA Philipps G., Gigot C.;
 RA Submitted (NOV-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the L18AE family of ribosomal proteins.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AC004077; AAC26708.1; -
 CC EMBL; AC004481; AAM14956.1; -
 CC EMBL; AY042803; AAK68743.1; -
 CC EMBL; Z18039; CAA79087.1; -
 CC PIR; T02335; T02335.
 CC InterPro; IPR002670; Ribosomal_L18ae.

DR Pfam: PF01775; Ribosomal_L18ae; 1.
KW Ribosomal protein.
FT CONFLICT 158 160 VRP -> LDH (IN REF. 3).
SQ SEQUENCE 178 AA; 21307 MW; 1F409C9B3CF27ED CRC64;

Query Match 41.6%; Score 42; DB 1; Length 178;
Best Local Similarity 45.0%; Pred. No. 10;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 VLPFPPKMKGLFSQAKISLF 20
DB 158 VRPPSKLTKTYKANKPNLF 177

RESULT 20
NK26 MOUSE
ID NK26 MOUSE STANDARD; PRT; 289 AA.
AC P43688;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Homeobox protein Nkx-2.6 (Homeobox protein Nkx-2 homolog F).
DE Nkx2-6 OR Nkx-2.6 OR Nkx2F.
GN Nkx2-6
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_TaxID=10090;
OX
RN
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=98213637; PubMed=9545560;
RA Biben C., Hatizistavrou T., Wang C.-C., Harvey R.P.;
RT "Expression of NK-2 class homeobox gene Nkx2-6 in foregut endoderm and heart.";
RT Mech. Dev. 73:125-127(1998).
RL
RN
RP SEQUENCE OF 123-214 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=94116443; PubMed=7904557;
RA Lints T.J., Parsons L.M., Hartley L., Lyons I., Harvey R.P.;
RT "Nkx-2.5: a novel murine homeobox gene expressed in early heart progenitor cells and their myogenic descendants.";
RT Development 119:419-431(1993).
CC
CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -1- TISSUE SPECIFICITY: Expressed in developing gut endoderm, cardiac progenitors and heart. Also expressed in neonatal brain.
CC -1- SIMILARITY: Belongs to the NK-2 homeobox family.
CC -1- SIMILARITY: Contains 1 homeobox domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF045150; AAC15674.1; --
DR HSSP; P23441; 1FTT.
DR TRANSFAC; T04342; --
DR MGD; MGI:97351; Nkx2-6.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF000046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SMC0389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.
FT DNA BIND 123 182 HOMEBOX.
SQ SEQUENCE 289 AA; 31597 MW; DBP8740A39CC0805 CRC64;

Query Match 40.6%; Score 41; DB 1; Length 289;
Best Local Similarity 53.8%; Pred. No. 10;
Matches 7; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKI 17
DB 262 PAKMRGILSQAMV 274

RESULT 22
AXOL RAT
ID AXOL RAT STANDARD; PRT; 1040 AA.
AC P22063;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)

Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPRKMKGLFSQAKI 17
DB 122 POKSRVLFSSQAV 135

RESULT 21
MCAL CRIGR
ID MCAL CRIGR STANDARD; PRT; 359 AA.
AC O54873;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Multisynthetase complex auxiliary component p43 [Contains;
DE Endothelial-monocyte activating polypeptide II (EMAP-II) (Small
DE Inducible cytokine subfamily E member 1)].
GN SCYB1.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98070438; PubMed=9405472;
RA Quevillon S., Agou F., Robinson J.-C., Mirande M.;
RT "The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine.";
RT J. Biol. Chem. 272:32573-32579(1997).
CC -1- SUBUNIT: Component of the multisynthetase complex which is comprised of a bifunctional glutamyl-prolyl-tRNA synthetase, the monospecific isoleucyl, leucyl, glutamyl, methionyl, lysyl, arginyl, and aspartyl-tRNA synthetases as well as three auxiliary proteins, p18, p48 and p43.
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF021800; AAB95207.1; --
DR InterPro; IPR008232; ENAP11.
DR InterPro; IPR008994; Nucleic acid_OB.
DR InterPro; IPR002547; tRNA_bind.
DR Pfam; PF01588; tRNA_bind; 1.
DR PIRSF; PIRSF005381; ENAP11; 1.
DR PROSITE; PS50886; TRBD; 1.
KW Protein biosynthesis; RNA-binding; tRNA-binding; Cytokine.
FT DOMAIN 198 299 TRNA-BINDING.
FT SEQUENCE 359 AA; 39601 MW; 4D868D1B65D72C23 CRC64;

Query Match 40.6%; Score 41; DB 1; Length 359;
Best Local Similarity 53.8%; Pred. No. 31;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKI 17
DB 262 PAKMRGILSQAMV 274

RESULT 22
AXOL RAT
ID AXOL RAT STANDARD; PRT; 1040 AA.
AC P22063;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)
Contactin 2 precursor (Axonin-1) (Axonal glycoprotein TAG-1)
(Transient axonal glycoprotein 1) (TAX-1).
CNTN2 OR TAX1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A., AND SEQUENCE OF 31-41.
RP TISSUE=Spinal cord;
RX MEDLINE=90199890; PubMed=2317872;
RA Fureley A.J., Morton S.B., Manalo D., Karagogeos D., Dodd J.,
RA Jessell T.M.;
RT "The axonal glycoprotein TAG-1 is an immunoglobulin superfamily
RT member with neurite outgrowth-promoting activity.";
RL Cell 61:157-170(1990).
CC -!- FUNCTION: May play a role in the initial growth and guidance of
CC axons. May be involved in cell adhesion.
CC -!- SUBCELLULAR LOCATION: Attached to the neuronal membrane by a GPI-
CC anchor and is also released from neurons.
CC -!- TISSUE SPECIFICITY: In neural tissues in embryos, and in adult
CC brain, spinal cord and cerebellum.
CC -!- DEVELOPMENTAL STAGE: Transiently expressed on a subset of axons
CC in the developing rat nervous system.
CC -!- SIMILARITY: Contains 6 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 4 fibronectin type III domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M31725; AAA42201.1; -;
DR PIR: A34695; A34695;
DR InterPro: IPR008957; FN III-like.
DR InterPro: IPR003961; FN-III.
DR InterPro: IPR003962; FNIII subd.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR Pfam: PF00041; fn3; 4.
DR Pfam: PF00047; Ig; 6.
DR PRINTS: PR00014; FNTYPEIII.
DR SMART: SM00060; FN3; 4.
DR SMART: SM00408; IGC2; 5.
DR PROSITE: PS00835; IG_LIKE; 6.
KW Immunoglobulin domain; Glycoprotein; Signal; GPI-anchor;
KW Cell adhesion; Repeat.
FT SIGNAL 1 30
FT CHAIN 31 1015
FT PROPER 1016 1040
FT DOMAIN 39 130
FT DOMAIN 135 224
FT DOMAIN 241 324
FT DOMAIN 329 413
FT DOMAIN 419 506
FT DOMAIN 511 605
FT DOMAIN 608 614
FT DOMAIN 613 708
FT DOMAIN 716 811
FT DOMAIN 818 910
FT DOMAIN 911 1005
FT SITE 796 798
FT CARBOHYD 78 78
FT CARBOHYD 200 200
FT CARBOHYD 206 206
FT CARBOHYD 463 463
FT CARBOHYD 479 479
FT CARBOHYD 500 500
FT CARBOHYD 527 527

FT CARBOHYD 777 777 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 920 920 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1040 AA; 113042 MW; 6E707EF6614CB4FB CRC64;
Query Match 40.6%; Score 41; DB 1; Length 1040;
Best Local Similarity 43.5%; Pred. No. 94;
Matches 10; Conservative 3; Mismatches 6; Indels 4; Gaps 1;
QY 1 VLPPPRMKG----LFSQAKISL 19
DB 910 VKPPRRPPGNISWTFSSLSL 932
RESULT 23
ID JMJ MOUSE STANDARD; PRT; 1234 AA.
AC Q62315;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Jumonji protein.
GN JMJ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Ola;
RX MEDLINE=95278734; PubMed=7758946;
RA Takeuchi T., Yamazaki Y., Katoh-Fukui Y., Tsuchiya R., Kondo S.,
RA Motoyama J., Higashinakagawa T.;
RT "Gene trap capture of a novel mouse gene, jumonji, required for
RT neural tube formation.";
RL Genes Dev. 9:1211-1222(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SUBCELLULAR LOCATION, AND FUNCTION.
RX MEDLINE=20374468; PubMed=10913339;
RA Toyoda M., Kojima M., Takeuchi T.;
RT "Jumonji is a nuclear protein that participates in the negative
RT regulation of cell growth.";
RL Biochem. Biophys. Res. Commun. 274:332-336(2000).
RN [4]
RP SUBCELLULAR LOCATION, AND FUNCTION.
RX MEDLINE=20269961; PubMed=10807864;
RA Lee Y., Song A.J., Baker R., Micales B., Conway S.J., Lyons G.E.;
RT "Jumonji, a nuclear protein that is necessary for normal heart

RT development.",
 RL Circ. Res. 86:932-938(2000).
 CC -!- FUNCTION: Required for neural tube formation. Essential for normal
 CC heart development and function. Participates in the negative
 CC regulation of cell proliferation signaling.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Brain.
 CC -!- SIMILARITY: Contains 1 ARID domain.
 CC -!- SIMILARITY: Contains 1 JMW domain.
 CC -!- SIMILARITY: Contains 1 JMW domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D31967; BAA06736.1; -
 CC EMBL; BC052444; AAH52444.1; -
 CC EMBL; BC060695; AAH60695.1; -
 CC F1; T30254; T30254.
 CC MGD; MGI:104813; JmJ.
 CC InterPro; IPR001606; ARID.
 CC InterPro; IPR003347; TF_JmJ.
 CC InterPro; IPR003349; TF_JmJ.
 CC InterPro; IPR004198; ZNF_C5HC2.
 CC Pfam; PF01388; ARID; 1.
 CC Pfam; PF02373; JmJC; 1.
 CC Pfam; PF02375; JmJN; 1.
 CC Pfam; PF02928; ZF-C5HC2; 1.
 CC SMART; SM00501; BRIGHT; 1.
 CC SMART; SM00558; JmJC; 1.
 CC SMART; SM00545; JmJN; 1.
 CC Developmental protein, Nuclear protein.
 FT DOMAIN 616 726
 ARID.
 SQ SEQUENCE 1234 AA; 137445 MW; B56E172C5E5745B5 CRC64;
 Query Match 40.6%; Score 41; DB 1; Length 1234;
 Best Local Similarity 53.8%; Pred. No. 1.1e-02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 4 PPRKMKGLFSQAK 16
 Db 450 PPRKMKGVAGNAE 462
 RESULT 24
 ZNTR_HAEIN STANDARD; PRT; 135 AA.
 AC P45277;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE HTH-type transcriptional regulator zntr homolog.
 ZNTR OR H11623.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Rd / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleisemann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 RL Rd".
 RL Science 269:496-512(1995).
 CC -!- FUNCTION: Transcriptional regulator (By similarity).
 CC -!- SIMILARITY: Contains 1 HTH merR-type DNA-binding domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U32835; AAC23268.1; -
 CC TIGR; H11623; -
 CC InterPro; IPR000551; HTH_MerR.
 CC Pfam; PF00376; merR; 1.
 CC PRINTS; PR00040; HTHMERR.
 CC SMART; SM00422; HTH_MERR; 1.
 CC PROSITE; PS00552; HTH_MERR_1; 1.
 CC PROSITE; PS00937; HTH_MERR_2; 1.
 CC Transcription regulation; DNA-binding; Complete proteome.
 FT DOMAIN 1 69
 HTH_MERR-TYPE.
 FT DNA_BIND 4 23
 H-T-H MOTIF (POTENTIAL).
 SQ SEQUENCE 135 AA; 15636 MW; CE07D848D0BCF9C CRC64;
 Query Match 39.6%; Score 40; DB 1; Length 135;
 Best Local Similarity 42.9%; Pred. No. 16;
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 1 VLPKPKMKGLFSQ 14
 Db 25 LIPPKRTSGNFRQ 38
 RESULT 25
 YOWI_CABEL STANDARD; PRT; 292 AA.
 ID YOWI_CABEL
 AC FJ0647;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein ZK643.1 in chromosome III.
 ZK643.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=92168156; PubMed=1538779;
 RA Sulston J., Du Z., Thomas K., Wilson R., Hillier L., Staden R.,
 RA Halloran N., Green P., Thierry-Mieg J., Qiu L., Dear S., Coulson A.,
 RA Craxton M., Durbin R., Berks M., Metzstein M., Hawkins T.,
 RA Ainscough R., Waterston R.
 RT "The C. elegans genome sequencing project: a beginning."
 RL Nature 356:37-41(1992).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z11126; CAAT7470.1; -
 CC F1; S23239; S23239.
 CC WormPep; ZK643.1; CE00440.
 CC InterPro; IPR000698; Arrestin.

Kao J., Fan Y., Haehnel I., Brett J., Greenberg S., Clauss M., Kayton M., Houck K., Kisiel W., Seljelid R., Burnier J., Stern D.; "A peptide derived from the amino terminus of endothelial-monocyte-activating polypeptide II modulates mononuclear and polymorphonuclear leukocyte functions, defines an apparently novel cellular interaction site, and induces an acute inflammatory response."; J. Biol. Chem. 269:9774-9782(1994).

-!- FUNCTION: ALTERS ENDOTHELIAL AND MONOCYTE FUNCTIONS, INDUCES THE MIGRATION OF MONOCYTES AND GRANULOCYTES, AND INDUCES AN INFLAMMATORY RESPONSE IN THE MOUSE FOOTPAD MODEL. EMAP II ELICITS A PHLOGENIC RESPONSE AND, POTENTIALLY, AUGMENTS THE EFFECTS OF THE OTHER TUMOR-DERIVED CYTOKINES.

-!- SUBUNIT: Monomer.

-!- SIMILARITY: Contains 1 tRNA-binding domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; U10118; AAA62203.1; -.
EMBL; BC002054; AAH02054.1; -.
PIR; A5053; A5053.
MGD; MGI:102774; Scyel.
InterPro; IPR008232; EMAPII.
InterPro; IPR008994; Nucleic acid_OB.
InterPro; IPR002547; tRNA_bind.
Pfam; PF01588; tRNA_bind; 1.
PIRSF; PIRSF005381; EMAPII; 1.
PROSITE; PS50886; TRBD; 1.
Protein biosynthesis; RNA-binding; tRNA-binding; Cytokine.
PROPEP 1 144
CHAIN 145 310
DOMAIN 149 250
SEQUENCE 310 AA; 33997 MW; A2F8FF52A33D03A0 CRC64;
QUERY MATCH 39.6%; Score 40; DB 1; Length 310;
Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 5 PRKMKGLFSQAKI 17
| | | : | | :
Db 213 PAKMGVLSQLMV 225

RESULT 27
MCAL HUMAN
ID MCAL HUMAN STANDARD; PRT; 312 AA.
AC Q12904; Q96CQ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Multisynthetase complex auxiliary component p43 [Contains:
DE Endothelial-monocyte activating polypeptide II (Small
DE inducible cytokine subfamily E member 1)].
DE SCYL OR EMAP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=95014290; PubMed=7929199;
RA Kao J., Houck K., Fan Y., Haehnel I., Libutti S.K., Kayton M.L., Grikscheit T., Chabot J., Nowyrod R., Greenberg S., Kuang W.J., Leung D.W., Hayward J.R., Kisiel W., Heath M., Brett J., Stern D.M.; "Characterization of a novel tumor-derived cytokine. Endothelial-monocyte activating polypeptide II."; J. Biol. Chem. 269:25106-25119(1994).

```

Query Match          39.6%; Score 40; DB 1; Length 310;
Best Local Similarity 53.8%; Pred. No. 38;
Matches 7; Conservative 3; Mismatches 0; Gaps 0;

QY      5 PRKMGGLFSQAKI 17
      | | | | |
DB      213 PAKMRGVLQAMV 225

RESULT 27
MCAL HUMAN
ID MCAL HUMAN STANDARD; PRT; 312 AA.
AC Q12904; Q96CQ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Multisynthetase complex auxiliary component p43 [Contains:
DE Endothelial-monocyte activating polypeptide II (EMAP-II) (Small
DE inducible cytokine subfamily E member 1)].
DE SCY1 OR EMAP2.
GN Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95014290; PubMed=7929199;
RA Kao J., Houck K., Fan Y., Haehnel I., Libutti S.K., Kayton M.L.,
RA Grikscheit T., Chabot J., Nowyrod R., Greenberg S., Kuang W.J.,
RA Leung D.W., Hayward J.R., Kisiel W., Heath M., Brett J., Stern D.M.;
RT "Characterization of a novel tumor-derived cytokine. Endothelial-
RT monocyte activating polypeptide II.";
RL J. Biol. Chem. 269:25106-25119(1994).

```

[2]
SEQUENCE FROM N.A.
TISSUE=pancreas;
MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
EL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SIMILARITY: Contains 1 tRNA-binding domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; U10117; AAA62202.1; -;
DR EMBL; BC014051; AAH14051.1; -;
DR PDB; 1E7Z; 06-FEB-01.
DR PDB; 1E7Z; 06-SEP-00.
DR PDB; 1FLO; 07-FEB-01.
DR Gene; HGNC:10648; SCYL1.
DR MIM; 603605; -;
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005125; P:cytokine activity; TAS.
DR GO; GO:0000049; F:tRNA binding; TAS.
DR GO; GO:0006418; P:amino acid activation; TAS.
DR GO; GO:0006935; P:chemotaxis; TAS.
DR GO; GO:0006954; P:inflammatory response; TAS.
DR InterPro; IPR008232; EMAP11.
DR InterPro; IPR008994; Nucleic acid OB.
DR Pfam; PF01588; tRNA_bind.1.
DR PIRSF; PIRSF005381; EMAP11.1.
DR PROSITE; PS00886; TRBD; 1.
KW protein biosynthesis; RNA-binding; tRNA-binding; Cytokine;
KW 3D-structure.
FT PROPEP 1 146
FT CHAIN 147 312
FT EMBL; ENDOTHELIAL-MONOCYTE ACTIVATING
FT POLYPEPTIDE II.
FT DOMAIN 151 252
FT TRNA-BINDING.
FT CONFLICT 79 79
FT A -> P (IN REF. 2).
SQ SEQUENCE 312 AA; 34356 MW; 946310A0216F7587 CRC64;
Query Match 39.6%; Score 40; DB 1; Length 312;
Best Local Similarity 53.8%; Pred. No. 39;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 5 PEKMGGLPSQAKI 17
DB 215 PAMRGVLVSQAVV 227
RESULT 28
NK24_HUMAN

DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein; Developmental protein.
FT DNA BIND 189 248 HOMEBOX.
FT DOMAIN 44 47 POLY-ALA.
FT DOMAIN 76 91 POLY-ALA.
FT DOMAIN 181 187 POLY-ALA.
FT DOMAIN 264 272 POLY-PRO.
SQ SEQUENCE 353 AA; 36011 MW; B6780C4E4020BED6 CRC64;

Query Match 39.6%; Score 40; DB 1; Length 353;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPRKMKGLFSQAKI 17
Db 188 PRKRRVLFSSQAV 201

RESULT 29
NK24 MOUSE STANDARD; PRT; 354 AA.
AC Q9EQM3; Q9EQM4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Homeobox protein Nkx-2.4 (Homeobox protein NKX2.4) (Homeobox protein
DN NK-2 homolog D).
GN NKX2-4 OR NKX2D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV, and C57BL/6;
RX MEDLINE=20279852; PubMed=10818213;
RA Wang C.-C., Brodnicki T., Copeland N.G., Jenkins N.A., Harvey R.P.;
RT "Conserved linkage of NK-2 homeobox gene pairs Nkx2-2/2-4 and
RL Nkx2-1/2-9 in mammals.";
RN Mamm. Genome 11:466-468(2000).
RN [2]
RP PARTIAL SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=92153416; PubMed=1346742;
RA Price M., Lazzaro D., Pohl T., Mattei M.-G., Ruether U., Olivo J.-C.,
RA Duboule D., Di Lauro R.;
RT "Regional expression of the homeobox gene Nkx-2.2 in the developing
RT mammalian forebrain.";
RL Neuron 8:241-255(1992).
CC -!- FUNCTION: Probable transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- TISSUE SPECIFICITY: In the embryo it is detected in the posterior
CC hypothalamus and later in the head. In the adult it is detected
CC only in testis.
CC -!- DEVELOPMENTAL STAGE: Expressed in a restricted region of the
CC posterior hypothalamus from E10. Detected in the head region from
CC E12.5 to E14.5. Expression is down-regulated by E15.5.
CC -!- SIMILARITY: Belongs to the NK-2 homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF202038; AAC35618.1; -.
DR EMBL; AF202039; AAC35619.1; -.
DR HSSP; P23441; 1FTT.
DR MGD; MGI:97349; Nkx2-4.
DR InterPro; IPR001356; Homeobox.

DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW DNA-binding; Homeobox; Nuclear protein; Developmental protein.
FT DNA BIND 188 247 HOMEBOX.
FT DOMAIN 76 91 POLY-ALA.
FT DOMAIN 132 136 POLY-ALA.
FT DOMAIN 263 272 POLY-PRO.
SQ SEQUENCE 354 AA; 36225 MW; A7A0F450DB15616C CRC64;

Query Match 39.6%; Score 40; DB 1; Length 354;
Best Local Similarity 57.1%; Pred. No. 44;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPRKMKGLFSQAKI 17
Db 187 PRKRRVLFSSQAV 200

RESULT 30
TTF1 CANFA STANDARD; PRT; 371 AA.
AC P43698;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thyroid transcription factor 1 (Thyroid nuclear factor 1) (TTF-1)
DE (Homeobox protein Nkx-2.1).
GN TTF1 OR TTF1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96034516; PubMed=7589789;
RX van Renterghem P.H.G., Dremier S., Vassar G., Christophe J.;
RT "Study of TTF-1 gene expression in dog thyrocytes in primary
RT culture.";
RL Mol. Cell. Endocrinol. 112:83-93(1995).
CC -!- FUNCTION: Transcription factor that binds and activates the
CC promoter of thyroid specific genes such as thyroglobulin,
CC thyroperoxidase, and thyrotropin receptor. Crucial in the
CC maintenance of the thyroid differentiation phenotype. May play a
CC role in lung development and surfactant homeostasis.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Thyroid, lung and CNS.
CC -!- SIMILARITY: Belongs to the NK-2 homeobox family.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X77910; CRA54868.1; -.
DR PIR; I46089; I46089.
DR HSSP; P23441; 1FTT.
DR TRANSFAC; T02098; -.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD00010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS50071; HOMEBOX 2; 1.
KW Transcription regulation; Activator; Homeobox; DNA-binding;

KW Nuclear protein. 220 HOMBBOX.
FT DNA BIND 161 220
FT DOMAIN 234 243 POLY-GLY.
FT DOMAIN 246 253 POLY-GLN.
FT DOMAIN 294 303 POLY-ALA.
SQ SEQUENCE 371 AA; 38539 MW; 3F16CE8E562604D7 CRC64;
Query Match 39.6%; Score 40; DB 1; Length 371;
Best Local Similarity 57.1%; Pred. No. 46;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPRKMKGLFSQAKI 17
D5 160 PPRKRVLFSAQV 173

Search completed: May 11, 2004, 13:38:05
Job time : 9.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:28:26 ; Search time 32 Seconds
(without alignments)
197.199 Million cell updates/sec

Title: US-09-171-432a-42
Perfect score: 101
Sequence: 1 VLPFRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

- SPTREMBL 25:*
- 1: sp archaea:*
 - 2: sp bacteria:*
 - 3: sp fungi:*
 - 4: sp human:*
 - 5: sp invertebrate:*
 - 6: sp mammal:*
 - 7: sp mhc:*
 - 8: sp organelle:*
 - 9: sp phage:*
 - 10: sp plant:*
 - 11: sp rodent:*
 - 12: sp virus:*
 - 13: sp vertebrate:*
 - 14: sp unclassified:*
 - 15: sp rvirus:*
 - 16: sp bacteriaph:*
 - 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
2	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
3	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
4	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
5	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
6	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
7	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
8	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
9	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
10	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
11	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
12	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
13	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
14	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
15	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a
16	101	100.0	94	12 Q9ENU9	Q9enu9 hepatitis a

90 Q9ENT4 hepatitis a
91 Q9ENV0 hepatitis a
92 Q9ENT7 hepatitis a
93 Q9ENT6 hepatitis a
94 Q86534 hepatitis a
95 Q9DWR5 hepatitis a
96 Q9W74 hepatitis a
97 Q71975 hepatitis a
98 Q71978 hepatitis a
99 Q8B8K2 hepatitis a
100 Q913V2 hepatitis a

ALIGNMENTS

RESULT 1
Q9ENT9 PRELIMINARY; PRT; 94 AA.
AC Q9ENT9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 32;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AB038298; BAB11836.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0430; PQ0430.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 2
Q9ENV9 PRELIMINARY; PRT; 94 AA.
AC Q9ENV9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 03;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.

DR EMBL: AB038278; BAB11816.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFDAB8C4BAC CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 3
Q9ENV5 PRELIMINARY; PRT; 94 AA.
AC Q9ENV5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 08;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AB038282; BAB11820.1; --
DR PIR: PQ0427; PQ0427.
DR PIR: PQ0428; PQ0428.
DR PIR: PQ0430; PQ0430.
DR PIR: PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF9BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 4
Q9ENV2 PRELIMINARY; PRT; 94 AA.
AC Q9ENV2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 28;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB038295; BAB11833.1; -.

DR PIR; PQ0427; PQ0428.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

FT NON_TER 1 94

FT NON_TER 94 94

SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.6e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20

|||||

Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 5

Q9ENV1

ID Q9ENV1 PRELIMINARY; PRT; 94 AA.

AC Q9ENV1;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OC NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Kantou 49;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,

RA Yasuoka A., Oka S.;

RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AB038306; BAB11844.1; -.

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

FT NON_TER 1 94

FT NON_TER 94 94

SQ SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.6e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20

|||||

Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 6

Q9ENV3

ID Q9ENV3 PRELIMINARY; PRT; 94 AA.

AC Q9ENV3;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OC NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Nagasaki 12n;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,

RA Yasuoka A., Oka S.;

RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AB038284; BAB11822.1; -.

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

FT NON_TER 1 94

FT NON_TER 94 94

SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE8A1A364BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.6e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20

|||||

Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 7

Q9ENV1

ID Q9ENV1 PRELIMINARY; PRT; 94 AA.

AC Q9ENV1;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OC NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Nagasaki 29;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,

RA Yasuoka A., Oka S.;

RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RL EMBL; AB038296; BAB11834.1; -.

DR PIR; PQ0427; PQ0427.

DR PIR; PQ0428; PQ0428.

DR PIR; PQ0430; PQ0430.

DR PIR; PQ0431; PQ0431.

FT NON_TER 1 94

FT NON_TER 94 94

SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;

Best Local Similarity 100.0%; Pred. No. 2.6e-09;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20

|||||

Db 74 VLPPPRKMKGLFSQAKISLF 93

RESULT 8

Q9ENV8

ID Q9ENV8 PRELIMINARY; PRT; 94 AA.

AC Q9ENV8;

DT 01-MAR-2001 (TREMELrel. 16, Created)

DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)

DE Polyprotein (Fragment).

OS Hepatitis A virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.

OC NCBI_TaxID=12092;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Nagasaki 22;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038285; BAB11827.1; -.
 DR PIR; PQ0427; PQ0428.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93
 RESULT 9
 Q9ENV2 PRELIMINARY; PRT; 94 AA.
 ID Q9ENV2
 AC Q9ENV2;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis A virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 15;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038285; BAB11823.1; -.
 DR PIR; PQ0427; PQ0428.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93
 RESULT 10
 Q9ENV5 PRELIMINARY; PRT; 94 AA.
 ID Q9ENV5
 AC Q9ENV5;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis A virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 15;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038285; BAB11823.1; -.
 DR PIR; PQ0427; PQ0428.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038285; BAB11827.1; -.
 DR PIR; PQ0427; PQ0428.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93
 RESULT 11
 Q9ENV1 PRELIMINARY; PRT; 94 AA.
 ID Q9ENV1
 AC Q9ENV1;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis A virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Hepatovirus.
 OC NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 18;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038286; BAB11824.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93
 RESULT 12
 Q9ENV9 PRELIMINARY; PRT; 94 AA.
 ID Q9ENV9
 AC Q9ENV9;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Polyprotein (fragment).
 OS Hepatitis A virus.
 OS NCBI_TaxID=12092;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nagasaki 18;
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
 RA Yasuoka A., Oka S.,
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB038286; BAB11824.1; -.
 DR PIR; PQ0427; PQ0427.
 DR PIR; PQ0428; PQ0428.
 DR PIR; PQ0430; PQ0430.
 DR PIR; PQ0431; PQ0431.
 DR PIR; PQ0431; PQ0431.
 FT NON_TER 1
 FT NON_TER 94
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;
 Query Match 100.0%; Score 101; DB 12; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.6e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLPPPRKMKGLFSQAKISLF 20
 Db 74 VLPPPRKMKGLFSQAKISLF 93

Query Match	100.0%;	Score 101;	DB 12;	Length 94;
Best Local Similarity	100.0%;	Pred. No. 2.6e-09;		
Matches	20;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;

QY 1 VLPPPRKMGLFSQAKISLF 20
|||
74 VT.PPPRKMGLFSQAKISLF 93
|||

RESULT 13

Q9ENV7	PRELIMINARY;	PRT;	94 AA.
ID	Q9ENV7		
AC	Q9ENV7;		
DT	01-MAR-2001 (TrEMBLrel. 16, Created)		
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Polyprotein (Fragment);		
OS	Hepatitis A virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;		
OC	Hepatovirus.		
OX	NCBI_TaxID=12092;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Nagaaki 06;		
RA	Iida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,		
RA	YasuoKa A., Oka S.;		
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)		
RT	infection: Prolonged HAV viremia and mild liver injury.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AB038280; BAB11818.1; -;		
DR	PIR; PQ0427; PQ0427.		
DR	PIR; PQ0428; PQ0428.		
DR	PIR; PQ0430; PQ0430.		
DR	PIR; PQ0431; PQ0431.		
FT	NON TER	1	
FT	NON TER	94	
FT	NON TER	94	
SO	SEQUENCE	94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;	

Query Match	Score 101;	DB 12;	Length 94;
Best Local Similarity	100.0%;		
Matches	20;		
Conservative	0;		
Mismatches	0;		
Indels	0;		
Gaps	0;		

QY 1 VLPPPRKMKGLFSQAKISLF 20
|||
74 VLPPPRKMKGLFSQAKISLF 93

RESULT 14

Q9ENV8	PRELIMINARY;	PRT;	94 AA.
ID	Q9ENV8		
AC	Q9ENV8;		
DT	01-MAR-2001	(TRENBLrel. 16, Created	
DT	01-MAR-2001	(TRENBLrel. 16, Last sequence update)	

DT	01-OCT-2003 (TREMBleref. 25, Last annotation update)
DE	Polyprotein (fragment).
OS	Hepatitis A virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage
OC	Hepatovirus.
OX	NCBI_TaxID=12092;
	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=Nagasaki 04;
RC	Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano
RA	Yasuoka A., Oka S.;
RA	"Influence of HIV-1 infection on acute hepatitis A
RT	infection: Prolonged HAV viremia and mild liver infa
RL	Submitted (F85-2001) to the EMBL/GenBank/DBJ data
RL	EMBL; AB038279; BAB11817.1; -.
DR	PIR; P00427; P00427.
DR	PIR; P00428; P00428.
DR	PIR; P00430; P00430.
DR	PIR; P00431; P00431.
FT	NON_TER 1
FT	NON_TER 94
FT	NON_TER 94
SC	SCENTENCE 94 AB. 10876 MW. 9AF8FE91BB8C4BAC CRC

```

Query Match      100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 VLPPPRKMGLEFSQAKISLF 20
|||
74 VLPPPRKMGLEFSQAKISLF 93
pb

RESIT, T 15

Q9ENV6	PRELIMINARY;	PRT;	94 AA.
IC	Q9ENV6		
AD	Q9ENV6;		
DT	01-MAR-2001	(TEMBLrel. 16, Created)	
DT	01-MAR-2001	(TEMBLrel. 16, Last sequence update)	
DT	01-OCT-2003	(TEMBLrel. 25, Last annotation update)	
DE	Polyprotein (Fragment).		
OS	Hepatitis A virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatocivirus		
OC	NCBI_taxid=12092;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=Nagasaki 07;		
RC	Ida S.;		
RA	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)		
RT	infection: Prolonged HAV viremia and mild liver injury.";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
DI			

```

FI      NON_ITER      54      54
SO      SEQUENCE      94      94  AA:  10876  MW:  9AF8BE91BB8C4BAC  CRC64:

```

Query Match	100.0%;	Score 101;	DB 12;	Length 94;
Best Local Similarity	100.0%;	Pred. No. 2.6e-09;		
Matches	20.	Conservative	0.	Mismatches 0;
				Indels 0;
				Gaps 0;

```

Qy      1 VLPPPRKMKGLFSQAKISLF 20
          |||||
nb      74 VTLPPIRKMKGLEFSQAKISLF 93

```

RESIT.T 16

Q9ENW0	PRELIMINARY;	PRT;	94 AA.
ID Q9ENW0			
AC Q9ENW0;			

```

DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
DE Hepatitis A virus.
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
CX [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=Nagasaki 02;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RT Yasuoka A., Oka S.;
RE "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RE infection: Prolonged HAV viremia and mild liver injury.";
RE Submitted (F82-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB038277; BAB11815.1; -.
RR PIR; PQ0427; PQ0427.
RR PIR; PQ0428; PQ0428.
FT 1
FT NON TER 1
FT NON TER 94
FT NON TER 94
SQ SEQUENCE 94 AA; 10903 MW; 9AF8BE85D3D57329C CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VLPPPRKMGIFSQAISLF 20
 |||||
Db 74 VLPPPRKMGIFSQAISLF 93

RESULT 17
Q9ENU6
ID Q9ENU6
PRELIMINARY:
PRT: 94 AA.

Q9ENU6;	
AC	01-MAR-2001 (TrEMBLrel. 16, Created)
CD	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Polyprotein (Fragment).
OS	Hepatitis A virus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC	Hepatovirus.
OX	NCBI_TaxID=12092;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=Nagasaki 24;
RA	Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA	Yasuka A., Oka S.;
RT	"Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL	infection: Prolonged HAV viremia and mild liver injury.",
RRL	Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AB038291; BAB11829.1; -
DR	PIR; PQ0427; PQ0427.
DR	PIR; PQ0428; PQ0428.
DR	PIR; PQ0430; PQ0430.
DR	PIR; PQ0431; PQ0431.
FT	NON TER 1
FT	NON TER 94
TSO	SEQUENCE 94 AA; 10876 MW; 9AFBBE1BB8C4BAC CRG64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRMKGLFSQAKISLF 20
|||||
74 VLPPRRMKGLFSQAKISLF 93

RESULT 18
Q9ENUJ4
DDP, TMTNADP, 04 12
DDT.

```

RESULT 20
Q9NT5S PRELIMINARY; PRT; 94 AA.
ID Q9NT5S
AC Q9NT5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
[1]_ RN
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=Kantou 42;
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuoka A., Oka S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038302; EABI1840.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT NON_TER 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;
Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2,6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VLPFRKMKGLFSQAKISLF 20
Db 74 VLPFRKMKGLFSQAKISLF 93

```

```

RESULT 21
Q9ENU7 PRELIMINARY; PRT; 94 AA.
Q9ENU7;
Q9ENU7 AC 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
DE Hepatitis A virus.
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12092;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Nagasaki 23;
RC Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,
RA Yasuka A., Oka S.;
RR "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RT infection: Prolonged HAV viremia and mild liver injury.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RR EMBL; AB038290; BAB1828.1; -.
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON_TER 1
FT 94
FT NON_TER 94
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 2.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMGFLFSQAKISLF 20
|||||

```

```
FT CHAIN 65 >115 P2A PROTEIN.
FT NON TER 115 115
SQ SEQUENCE 115 AA; 13344 MW; 16DFB8E7EEBA77B0 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 115;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 96 VLPPrKMKGLFSQAKISLF 115
|||||

RESULT 24
O92941 PRELIMINARY; PRT; 116 AA.
AC O92941;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-1;
RX MEDLINE=97442186; PubMed=9298735;
RA Grinde B., Stene-Johansen K., Sharma B., Hoel T., Jensenius M.,
Skaug K.;
RT "Characterisation of an epidemic of hepatitis A virus involving
intra-venous drug abusers--infection by needle sharing?";
RL J. Med. Virol. 53:69-75(1997).
DR EMBL; AF050223; AAC05516.1; -
DR EMBL; AF050223; AAC05524.1; -
DR EMBL; AF050223; AAC05522.1; -
DR EMBL; AF050230; AAC05523.1; -
DR PIR; PQ0427; PQ0428.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0431; PQ0431.
KW Polyprotein.
FT NON TER 1 116
FT NON TER 116 116
SQ SEQUENCE 116 AA; 13479 MW; 4726C6A1E7EEBA64 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 96 VLPPrKMKGLFSQAKISLF 115
|||||

RESULT 25
O9W7X7 PRELIMINARY; PRT; 116 AA.
AC O9W7X7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-15;
RX MEDLINE=97442186; PubMed=9298735;
RA Grinde B., Stene-Johansen K., Sharma B., Hoel T., Jensenius M.,
Skaug K.;
RT "Characterisation of an epidemic of hepatitis A virus involving
intra-venous drug abusers--infection by needle sharing?";
RL J. Med. Virol. 53:69-75(1997).
DR EMBL; AF050226; AAC05519.1; -
DR EMBL; AF050225; AAC05518.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
KW Polyprotein.
FT NON TER 1 116
FT NON TER 116 116
SQ SEQUENCE 116 AA; 13507 MW; 8E26DFB8E7EEBA77 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 96 VLPPrKMKGLFSQAKISLF 115
|||||
```

```
RT "Characterisation of an epidemic of hepatitis A virus involving
intra-venous drug abusers--infection by needle sharing?";
RL J. Med. Virol. 53:69-75(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-15;
RA Stene-Johansen K., Skaug K., Blystad H., Grinde B.;
RT "An unique hepatitis A virus strain caused an epidemic in Norway
associated with intravenous drug abuse.";
RL Scand. J. Infect. Dis. 0:0-0(1998).
DR EMBL; AF050236; AAC05529.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
FT NON TER 1 116
FT NON TER 116 116
SQ SEQUENCE 116 AA; 13514 MW; 3E2AAE639A586106 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 96 VLPPrKMKGLFSQAKISLF 115
|||||

RESULT 26
O9W7S7 PRELIMINARY; PRT; 116 AA.
AC O9W7S7;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-4, and NOR-3;
RX MEDLINE=97442186; PubMed=9298735;
RA Grinde B., Stene-Johansen K., Sharma B., Hoel T., Jensenius M.,
Skaug K.;
RT "Characterisation of an epidemic of hepatitis A virus involving
intra-venous drug abusers--infection by needle sharing?";
RL J. Med. Virol. 53:69-75(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOR-4, and NOR-3;
RA Stene-Johansen K., Skaug K., Blystad H., Grinde B.;
RT "An unique hepatitis A virus strain caused an epidemic in Norway
associated with intravenous drug abuse.";
RL Scand. J. Infect. Dis. 0:0-0(1998).
DR EMBL; AF050226; AAC05519.1; -
DR EMBL; AF050225; AAC05518.1; -
DR PIR; PQ0427; PQ0427.
DR PIR; PQ0428; PQ0428.
DR PIR; PQ0430; PQ0430.
DR PIR; PQ0431; PQ0431.
KW Polyprotein.
FT NON TER 1 116
FT NON TER 116 116
SQ SEQUENCE 116 AA; 13507 MW; 8E26DFB8E7EEBA77 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 116;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPrKMKGLFSQAKISLF 20
DB 96 VLPPrKMKGLFSQAKISLF 115
|||||
```


Query Match	100.0%	Score 101;	DB 12;	Length 116;
Best Local Similarity	100.0%	Pred. No. 3.2e-09;		
Matches	20.	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;

Db 95 VLPPRRKMGFLFSQAKISLF 114

Search completed: May 11, 2004, 13:40:25
Job time : 33 secs

Run on: May 11, 2004, 13:34:17 ; Search time 13.5 Seconds

(without alignment)
76.483 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101
Sequence: 1 VLPPPRKMGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 rcs

Total number of hits satisfying chosen parameters:

Minimum DB seg length: 0

Maximum DB seq	length: 200000000
Maximum DB seq	length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.per.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.per:*

```
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pdb: *
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pdb: *
```

4: /cgiz_6/prodata/2/1aa/05 COMB.per:
5: /cgm2_6/prodata/2/1aa/PCITUS COMB.per:*

```
5: /cgn2_6/ptodata/2/iaa/cross_section.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*
```

1. *What is the purpose of this study?*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	101	100.0	1091	6	5516630-2	Patent No. 5516630
2	101	100.0	2227	3	US-08-475-886-2	Sequence 2, Appli
3	101	100.0	2227	3	US-08-475-886-4	Sequence 4, Appli
4	101	100.0	2227	3	US-08-475-886-6	Sequence 6, Appli
5	101	100.0	2227	3	US-08-397-232-2	Sequence 2, Appli
6	101	100.0	2227	3	US-08-397-232-4	Sequence 4, Appli
7	101	100.0	2227	3	US-09-171-387-2	Sequence 2, Appli
8	101	100.0	2227	4	US-09-653-499-2	Sequence 2, Appli
9	101	100.0	2227	4	US-09-653-499-4	Sequence 4, Appli
10	101	100.0	2227	4	US-09-653-499-6	Sequence 6, Appli
11	101	100.0	2227	4	US-10-104-966-12	Sequence 12, Appli
12	101	100.0	2227	4	US-10-135-988-2	Sequence 2, Appli
13	101	100.0	2227	4	US-10-135-988-4	Sequence 4, Appli
14	101	100.0	2227	4	US-10-135-988-6	Sequence 6, Appli
15	84	83.2	839	1	US-08-087-016-2	Sequence 2, Appli
16	52.5	52.0	1025	2	US-08-304-309-2	Sequence 2, Appli
17	52.5	52.0	1025	3	US-08-991-942-2	Sequence 2, Appli
18	52.5	52.0	1025	3	US-09-138-103-2	Sequence 2, Appli
19	52.5	52.0	1025	5	PCI-U995-04567-4	Sequence 4, Appli
20	45.5	45.0	1025	5	PCI-U995-04567-2	Sequence 2, Appli
21	45	44.6	23	1	US-07-800-3648-4	Sequence 4, Appli
22	45	44.6	23	5	PCI-U991-03388-4	Sequence 4, Appli
23	44	43.6	454	4	US-09-134-001C-4438	Sequence 4438, Ap
24	43.5	43.1	97	4	US-09-205-258-855	Sequence 855, App
25	42	41.6	345	4	US-09-252-991A-30726	Sequence 30726, A
26	42	41.6	1197	4	US-09-252-991A-30833	Sequence 30833, A
27	41.5	41.1	3330	4	US-09-252-991A-23794	Sequence 23794, A

ALIGNMENTS

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTEIN, STEPHEN M.; PURCELL, ROBERT H.; RACANELLO, VINCENT R.; BAROUDY, BAIGIE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/789,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match 100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLP 20
Db 1060 VLPPPRKMKGLFSQAKISLP 1079

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLP 20
Db 823 VLPPPRKMKGLFSQAKISLP 842

RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A

; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLP 20
Db 823 VLPPPRKMKGLFSQAKISLP 842

RESULT 4
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLP 20
Db 823 VLPPPRKMKGLFSQAKISLP 842

RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 7
US-09-171-387-2
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/171,387
FILING DATE: 24-Mar-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Reiler
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-4229US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-171-387-2

Query Match 100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
|||
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 8
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLP... 20
DB 823 VLP... 842

RESULT 9
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLP... 20
DB 823 VLP... 842

RESULT 10
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLP... 20
DB 823 VLP... 842

RESULT 11
US-10-104-966-12
; Sequence 12, Application US/10104966
; Patent No. 6680059
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLP... 20
DB 823 VLP... 842

RESULT 12
US-10-135-988-2
; Sequence 2, Application US/10135988
; Patent No. 6680060
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLP... 20
DB 823 VLP... 842

GENERAL INFORMATION:
APPLICANT: NAINAN, OMANA V.
APPLICANT: MARGOLIS, HAROLD S.
APPLICANT: ROBERTSON, BETTY H.
APPLICANT: BRINTON, MARGO H.
APPLICANT: EBERT, JAMES W.
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
STREET: 1615 L Street N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/087,016
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,828
FILING DATE: 03-APR-1991
ATTORNEY/AGENT INFORMATION:
NAME: SCOTT, WARSON I.
REGISTRATION NUMBER: 26,581
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-087-016-2

Query Match 83.2%; Score 84; DB 1; Length 839;
Best Local Similarity 83.3%; Pred. No. 6.6e-06;
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 18
Db 822 ILPPPRKLGKGLFSQSKIS 839

RESULT 16
US-08-304-309-2
Sequence 2, Application US/08304309
Patent No. 5856454
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 13
US-10-135-988-4
Sequence 4, Application US/10135988
Patent No. 6680060
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 14
US-10-135-988-6
Sequence 6, Application US/10135988
Patent No. 6680060
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patent in Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match 100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 15
US-08-087-016-2
Sequence 2, Application US/08087016
Patent No. 5430135

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/304,309
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1025 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-304-309-2

Query Match 52.0%; Score 52.5; DB 2; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.7;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPEKMEAYS-AKIALF 193

RESULT 17
US-08-991-942-2
Sequence 2, Application US/08991942
Patent No. 6015673
GENERAL INFORMATION:
APPLICANT: GONZALEZ, Frank J.
APPLICANT: FERNANDEZ-SALGUERO, Pedro
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: Steuart Street Tower, One Market Plaza
CITY: San Francisco
STATE: California
COUNTRY: US
ZIP: 94105-1493
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/991,942
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/304,309
FILING DATE: 09-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 15280-210
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 543-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1025 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-991-942-2

Query Match 52.0%; Score 52.5; DB 3; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.7;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPEKMEAYS-AKIALF 193

RESULT 18
US-09-138-103-2
Sequence 2, Application US/09138103A
Patent No. 6232448
GENERAL INFORMATION:
APPLICANT: Yoshikubo, Takashi
APPLICANT: Hasegawa, Masami
TITLE OF INVENTION: Dihydropyrimidine Dehydrogenase
TITLE OF INVENTION: Dihydropyrimidine Dehydrogenase
FILE REFERENCE: 09/138,103 Yoshikubo, et al.
CURRENT APPLICATION NUMBER: US/09/138,103A
CURRENT FILING DATE: 1998-08-21
EARLIER APPLICATION NUMBER: 97114630.3
EARLIER FILING DATE: 1997-08-22
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1025
TYPE: PRT
ORGANISM: Homo sapiens
US-09-138-103-2

Query Match 52.0%; Score 52.5; DB 3; Length 1025;
Best Local Similarity 63.2%; Pred. No. 1.7;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPEKMEAYS-AKIALF 193

RESULT 19
PCT-US95-04567-4
Sequence 4, Application PC/TUS9504567
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version
SOFTWARE: #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04567
FILING DATE: CONCURRENTLY HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/227,357
FILING DATE: 13-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Wilson, Mark B.
REGISTRATION NUMBER: 37,259
REFERENCE/DOCKET NUMBER: UOAB025P--

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (512) 418-3000
 TELEFAX: (713) 789-2679
 TELEX: 79-0924
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1025 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT-US95-04567-4

Query Match 52.0%; Score 52.5; DB 5; Length 1025;
 Best Local Similarity 63.2%; Pred. No. 1.7;
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGLFSQAKISL 20
 DB 176 LPPPRKMKGLFSQAKISL 193

RESULT 20
 PCT-US95-04567-2
 ; Sequence 2, Application PC/TUS9504567
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: United States of America
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
 ; SOFTWARE: Patent In Release #1.0, Version
 ; SOFTWARE: #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/04567
 ; FILING DATE: CONCURRENTLY HERewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/227,357
 ; FILING DATE: 13-APR-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Wilson, Mark B.
 ; REGISTRATION NUMBER: 37,259
 ; REFERENCE/DOCKET NUMBER: UOAB025P--
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (512) 418-3000
 ; TELEFAX: (713) 789-2679
 ; TELEX: 79-0924
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1025 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-04567-2

Query Match 45.0%; Score 45.5; DB 5; Length 1025;
 Best Local Similarity 61.1%; Pred. No. 26;
 Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGLFSQAKISL 19
 DB 176 LPPPRKMKGLFSQAKISL 192

RESULT 21
 US-07-800-364B-4
 ; Sequence 4, Application US/07800364B
 ; Patent No. 5688678
 ; GENERAL INFORMATION:
 ; APPLICANT: Hewick, Rodney M.
 ; APPLICANT: Wang, Jack H.
 ; APPLICANT: Wozney, John H.
 ; APPLICANT: Celeste, Anthony J.
 ; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/800,364B
 FILING DATE: 26-NOV-1991
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Kapinos, Ellen J.
 REGISTRATION NUMBER: 32,245
 REFERENCE/DOCKET NUMBER: GI 5182A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-876-1170
 TELEFAX: 617-876-5851
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 23 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Bos taurus
 TISSUE TYPE: Bone
 US-07-800-364B-4

Query Match 44.6%; Score 45; DB 1; Length 23;
 Best Local Similarity 63.6%; Pred. No. 0.51;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPPPRKMKGLF 12
 DB 4 LPPPRKMKGLF 14

RESULT 22
 PCT-US91-03388-4
 ; Sequence 4, Application PC/TUS9103388
 ; GENERAL INFORMATION:
 ; APPLICANT: Hewick, Rodney M.
 ; APPLICANT: Wang, Jack H.
 ; TITLE OF INVENTION: Bone and Cartilage Inductive Proteins
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
 ; STREET: 87 CambridgePark Drive
 ; CITY: Cambridge
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02140

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US91/03388
; FILING DATE: 19910515
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,245
; REFERENCE/DOCKET NUMBER: G15182X-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Bos taurus
; TISSUE TYPE: Bone
; PCT-US91-03388-4

Query Match 44.6%; Score 45; DB 5; Length 23;
Best Local Similarity 63.6%; Pred. No. 0.51;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LPPPKMKGLF 12
|||||:|:|
Db 4 LPPPNKLPGIF 14

RESULT 23
US-09-134-001C-4438
; Sequence 4438, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4438
; LENGTH: 454
; TYPE: PPT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4438

Query Match 43.6%; Score 44; DB 4; Length 454;
Best Local Similarity 53.3%; Pred. No. 19;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 LPPPKMKGLFSQAK 16
|||:|:|:|
Db 179 IPFPDKYKGMFEQAK 193

RESULT 24
US-09-205-258-855
; Sequence 855, Application US/09205258
; Patent No. 6525174
```

```
; GENERAL INFORMATION:
; APPLICANT: Young et al
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007F1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-855

Query Match 43.1%; Score 43.5; DB 4; Length 97;
Best Local Similarity 47.4%; Pred. No. 4.3;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

Qy 2 LPPPRKMKGLFSQAKISLF 20
Db 52 VPPPLKMPGVLA-IRLSLF 69

RESULT 25
US-09-252-991A-30726
; Sequence 30726, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30726
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30726

Query Match 41.6%; Score 42; DB 4; Length 345;
Best Local Similarity 50.0%; Pred. No. 31;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PPRKMKGLFSQAKISL 19
Db 12 PARETRGLTSQKRVSL 27

RESULT 26
US-09-252-991A-30833
; Sequence 30833, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30833
; LENGTH: 1197
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1028)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-30833

Query Match 41.6%; Score 42; DB 4; Length 1197;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VLPPPRKMKGLFSQAK 16
Db 1012 VPPPRARKELRSQAK 1027

RESULT 27
US-09-252-991A-23794
; Sequence 23794, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23794
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23794

Query Match 41.1%; Score 41.5; DB 4; Length 330;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 10; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Qy 2 LPPPRKMKGLFSQ-AKISLF 20
Db 128 IPASSRFRKGFSTQAEIQLF 147

RESULT 28
US-09-621-976-7008
; Sequence 7008, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7008
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7008

; MOLECULE TYPE: protein
US-08-444-733-106

Query Match 40.6%; Score 41; DB 2; Length 108;
Best Local Similarity 46.7%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 PPRKMKGLFSQAKI 17
||| ||| : :
Db 57 PPRVPAKGVPEAPL 71

Search completed: May 11, 2004, 13:42:23
Job time : 15.5 secs

OM protein - protein search, using sw model

Run on: May 11, 2004, 13:37:37 ; Search time 34.25 Seconds
(without alignments)
162.083 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1140673 seqs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US10F_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB ID	Description
	Score	Match	Length		
1	101	100.0	20	10	US-09-171-432A-42
2	101	100.0	352	14	US-10-272-459-45
3	101	100.0	980	14	US-10-272-459-41
4	101	100.0	2227	9	US-09-929-955-12
5	101	100.0	2227	13	US-10-104-966-12
6	101	100.0	2227	13	US-10-135-988-2
7	101	100.0	2227	13	US-10-135-988-4
8	101	100.0	2227	13	US-10-135-988-6
9	70	69.3	836	14	US-10-272-459-40
10	65	64.4	25	10	US-09-171-432A-47
11	52.5	52.0	1025	9	US-09-854-086-2
12	51	50.5	84	12	US-10-424-599-242191
13	48	47.5	187	12	US-10-424-599-264591
14	48	47.5	524	15	US-10-320-797-3221
15	47	46.5	650	12	US-10-424-599-280225

16	46	45.5	73	12	US-10-424-599-196137	Sequence 196137,
17	46	45.5	158	12	US-10-424-599-265499	Sequence 265499,
18	45	44.6	171	15	US-10-094-749-2193	Sequence 2193, Ap
19	45	44.6	2827	12	US-10-424-599-256710	Sequence 256710,
20	45	44.6	2834	12	US-10-424-599-256711	Sequence 256711,
21	44	43.6	151	12	US-10-424-599-200960	Sequence 200960,
22	44	43.6	178	12	US-10-424-599-198250	Sequence 198250,
23	44	43.6	408	12	US-10-087-684-76	Sequence 76, Appl
24	44	43.6	408	12	US-10-218-779-76	Sequence 76, Appl
25	44	43.6	457	12	US-10-425-114-4364	Sequence 4364, A
26	44	43.6	490	12	US-10-425-114-43063	Sequence 43063, A
27	44	43.6	522	15	US-10-104-047-3665	Sequence 3665, Ap
28	44	43.6	673	14	US-10-176-306-2	Sequence 2, Appli
29	43.5	43.1	62	12	US-10-424-599-212719	Sequence 212719,
30	43.5	43.1	97	10	US-09-933-767-855	Sequence 855, App
31	43.5	43.1	97	12	US-10-004-860-855	Sequence 855, App
32	43.5	43.1	97	14	US-10-023-282-855	Sequence 855, App
33	43	42.6	178	12	US-10-424-599-192572	Sequence 192572,
34	43	42.6	178	12	US-10-424-599-192574	Sequence 192574,
35	42.5	42.1	80	12	US-10-424-599-261181	Sequence 261181,
36	42.5	42.1	103	12	US-10-424-599-284569	Sequence 284569,
37	42	41.6	20	10	US-09-171-432A-43	Sequence 43, Appl
38	42	41.6	53	12	US-10-424-599-263190	Sequence 263190,
39	42	41.6	78	12	US-10-424-599-165621	Sequence 165621,
40	42	41.6	79	14	US-10-106-598-4310	Sequence 4310, Ap
41	42	41.6	83	12	US-10-424-599-153362	Sequence 153362,
42	42	41.6	113	11	US-08-854-08A-4034	Sequence 4034, Ap
43	42	41.6	178	12	US-10-424-599-151916	Sequence 151916,
44	42	41.6	255	12	US-10-424-599-174296	Sequence 174296,
45	42	41.6	270	12	US-10-282-122A-67337	Sequence 67337, A
46	42	41.6	449	14	US-10-156-761-13588	Sequence 13588, A
47	42	41.6	669	12	US-10-282-122A-72423	Sequence 72423, A
48	42	41.6	847	12	US-10-282-511-8	Sequence 8, Appli
49	42	41.6	2053	13	US-10-017-216-2	Sequence 2, Appli
50	42	41.6	2053	14	US-10-325-430-12	Sequence 12, Appl
51	41.5	41.1	49	12	US-10-424-599-283033	Sequence 283033,
52	41.5	41.1	76	12	US-10-424-599-185254	Sequence 185254,
53	41	40.6	60	12	US-10-424-599-281008	Sequence 281008,
54	41	40.6	86	12	US-10-424-599-278781	Sequence 278781,
55	41	40.6	113	15	US-10-131-410-89	Sequence 89, Appl
56	41	40.6	156	12	US-10-424-599-159379	Sequence 159379,
57	41	40.6	178	12	US-10-424-599-151918	Sequence 151918,
58	41	40.6	194	12	US-10-424-599-221615	Sequence 221615,
59	41	40.6	306	12	US-10-424-599-179002	Sequence 179002,
60	41	40.6	2842	14	US-10-224-399A-3479	Sequence 3479, Ap
61	41	40.6	2905	8	US-08-424-550B-401	Sequence 401, App
62	41	40.6	2910	10	US-09-828-498-2	Sequence 2, Appli
63	40.5	40.1	86	12	US-10-424-599-165143	Sequence 165143,
64	40.5	40.1	105	12	US-10-424-599-259408	Sequence 259408,
65	40	39.6	64	12	US-10-424-599-163834	Sequence 163834,
66	40	39.6	95	12	US-10-424-599-193808	Sequence 193808,
67	40	39.6	101	15	US-10-264-049-4089	Sequence 4089, Ap
68	40	39.6	113	12	US-10-424-599-147825	Sequence 147825,
69	40	39.6	137	12	US-10-424-599-220719	Sequence 220719,
70	40	39.6	166	9	US-09-851-026-11	Sequence 11, Appl
71	40	39.6	166	10	US-09-910-169-3	Sequence 3, Appli
72	40	39.6	178	12	US-10-424-599-281376	Sequence 281376,
73	40	39.6	182	12	US-10-424-599-194866	Sequence 194866,
74	40	39.6	183	12	US-10-424-599-181763	Sequence 181763,
75	40	39.6	272	10	US-09-934-455-140	Sequence 140, App
76	40	39.6	272	15	US-10-325-068-200	Sequence 200, App
77	40	39.6	272	15	US-10-374-780A-2428	Sequence 2428, Ap
78	40	39.6	310	9	US-09-851-026-35	Sequence 35, Appl
79	40	39.6	312	9	US-09-851-026-36	Sequence 36, Appl
80	40	39.6	318	9	US-09-947-971-2	Sequence 2, Appli
81	40	39.6	320	12	US-09-925-298-472	Sequence 472, App
82	40	39.6	320	14	US-10-102-806-472	Sequence 472, App
83	40	39.6	354	12	US-10-425-114-65632	Sequence 65632, A
84	40	39.6	371	9	US-09-320-337-76	Sequence 76, Appl
85	40	39.6	385	12	US-10-424-599-237737	Sequence 237737,
86	40	39.6	401	15	US-10-236-031B-26	Sequence 26, Appl
87	40	39.6	406	12	US-10-282-122A-73403	Sequence 73403, A
88	40	39.6	448	12	US-10-282-122A-58700	Sequence 58700, A

89 40 39.6 453 12 US-10-282-122A-52293 Sequence 52293, A
90 40 39.6 466 12 US-10-425-114-37460 Sequence 37460, A
91 40 39.6 524 12 US-10-072-012-290 Sequence 290, App
92 40 39.6 562 12 US-10-425-114-70255 Sequence 70255, A
93 40 39.6 562 15 US-10-369-493-21428 Sequence 21428, A
94 40 39.6 666 9 US-09-815-242-4933 Sequence 4933, Ap
95 40 39.6 669 12 US-10-282-122A-57809 Sequence 57809, A
96 40 39.6 677 9 US-09-815-242-10663 Sequence 10663, A
97 40 39.6 677 12 US-10-282-122A-42475 Sequence 42475, A
98 40 39.6 679 12 US-10-282-122A-48489 Sequence 48489, A
99 40 39.6 851 10 US-09-991-936-1915 Sequence 1915, Ap
100 40 39.6 2623 12 US-10-412-699B-1519 Sequence 1519, Ap

ALIGNMENTS

RESULT 1
US-09-171-432A-42
; Sequence 42, Application US/09171432A
; Publication No. US20030187194A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Khudyakov, Yury E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; FILE OF INVENTION: Hepatitis A Virus Polypeptide
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..20
; OTHER INFORMATION: /label=YK-1318

US-09-171-432A-42
Query Match 100.0%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 VLPPIPRKMKGLFSQAKISLF 20
Db 1 VLPPIPRKMKGLFSQAKISLF 20

RESULT 2

US-10-272-459-45
; Sequence 45, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: OF 38.8 kDa
US-10-272-459-45
Query Match 100.0%; Score 101; DB 14; Length 352;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLPPIPRKMKGLFSQAKISLF 20
Db 195 VLPPIPRKMKGLFSQAKISLF 214

RESULT 3

US-10-272-459-41
; Sequence 41, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICHUANES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; FILE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 115.5 kDa
US-10-272-459-41
Query Match 100.0%; Score 101; DB 14; Length 980;
Best Local Similarity 100.0%; Pred. No. 7.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VLPPIPRKMKGLFSQAKISLF 20
Db 823 VLPPIPRKMKGLFSQAKISLF 842

RESULT 4

US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; TITLE OF INVENTION: METHODS OF USE THEREOF

FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 7
US-10-135-988-4
Sequence 4, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-10-135-988-4

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 8
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US3
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18

FILE REFERENCE: TRIPEP.23AUS2
CURRENT APPLICATION NUMBER: US/09/929,955
CURRENT FILING DATE: 2001-08-15
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
PRIOR APPLICATION NUMBER: 60/225,767
PRIOR FILING DATE: 2000-08-17
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 2227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12

Query Match 100.0%; Score 101; DB 9; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5
US-10-104-966-12
Sequence 12, Application US/10104966
Publication No. US20020155124A1
GENERAL INFORMATION:
APPLICANT: Matti Sallberg
APPLICANT: Catharina Hultgren
TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: TRIPEP.23AUSC1
CURRENT APPLICATION NUMBER: US/10/104,966
CURRENT FILING DATE: 2002-03-22
PRIOR APPLICATION NUMBER: 09/705,547
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: 60/229,175
PRIOR FILING DATE: 2000-08-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12
LENGTH: 2227
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match 100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 6
US-10-135-988-2
Sequence 2, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES


```

; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6

Query Match      100.0%; Score 101; DB 13; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 9
US-10-272-459-40
; Sequence 40, Application US/10272459
; Publication No. US20030124517A1
; GENERAL INFORMATION:
; APPLICANT: PICRUANTES, Sergio
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 40
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
; OTHER INFORMATION: of 94 kDa
US-10-272-459-40

Query Match      69.3%; Score 70; DB 14; Length 836;
Best Local Similarity 92.9%; Pred. No. 0.037;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQ 14
Db 823 VLPPPRKMKGLFSQ 836

RESULT 10
US-09-171-432A-47
; Sequence 47, Application US/09171432A
; Publication No. US20030187184A1
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; APPLICANT: Knudsyakov, Yuri E.
; TITLE OF INVENTION: Antigenically Reactive Regions of the
; TITLE OF INVENTION: Hepatitis A Virus Polyprotein
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESS: Kilpatrick Stockton LLP
; STREET: 3424 Peachtree Road, N.E.
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30326
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

; PRIOR APPLICATION NUMBER: US/09/171.432A
; FILING DATE: 23-NOV-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greene, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0231US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 949-2400
; TELEFAX: (404) 949-2499
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label= YK-1665
US-09-171-432A-47

Query Match      64.4%; Score 65; DB 10; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0055;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLF 12
Db 14 VLPPPRKMKGLF 25

RESULT 11
US-09-854-886-2
; Sequence 2, Application US/09854886
; Patent No. US20020072080A1
; GENERAL INFORMATION:
; APPLICANT: Yoshikubo, Takashi
; APPLICANT: Hasegawa, Masami
; TITLE OF INVENTION: Immunological Material and Methods for Detecting
; TITLE OF INVENTION: Dihydropyrimidine Dehydrogenase
; FILE REFERENCE: 100554-32887
; CURRENT APPLICATION NUMBER: US/09/854,886
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 09/138,103
; PRIOR FILING DATE: 1998-08-21
; PRIOR APPLICATION NUMBER: 97114630.3
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-854-886-2

Query Match      52.0%; Score 52.5; DB 9; Length 1025;
Best Local Similarity 63.2%; Pred. No. 22;
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPPRKMKGLFSQAKISLF 20
Db 176 LPPPRKMKSEAYS-AKIALF 193

RESULT 12
US-10-424-599-242191
; Sequence 242191, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 242191
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_60726C.1.pep
US-10-424-599-242191

Query Match          50.5%; Score 51; DB 12; Length 84;
Best Local Similarity 66.7%; Pred. No. 2.8;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 PPPRKMKGFLFSQAKI 17
   |||:|||||
Db 59 PFPKXKGFSPRKI 73

RESULT 13
US-10-424-599-264591
; Sequence 264591, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 264591
; LENGTH: 187
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_80946C.1.pep
US-10-424-599-264591

Query Match          47.5%; Score 48; DB 12; Length 187;
Best Local Similarity 50.0%; Pred. No. 19;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 3 PPPRKMKGFLFSQAKISLF 20
   |||:|||||
Db 122 PPRGFVGLFSRPTGFF 139

RESULT 14
US-10-320-797-3221
; Sequence 3221, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Broshtkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
```

```
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3221
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
US-10-320-797-3221

Query Match          47.5%; Score 48; DB 15; Length 524;
Best Local Similarity 43.8%; Pred. No. 55;
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 LPPPRKMKGFLFSQAKI 17
   |||:|||||
Db 5 LPPPKRQKSAYSQSQM 20

RESULT 15
US-10-424-599-280225
; Sequence 280225, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 280225
; LENGTH: 650
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_95066C.1.pep
US-10-424-599-280225

Query Match          46.5%; Score 47; DB 12; Length 650;
Best Local Similarity 62.8%; Pred. No. 97;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 LPPPRKMKGFLFSQAKI 17
   |||:|||||
Db 85 LPDPFKRKGRTQKKI 100

RESULT 16
US-10-424-599-196137
; Sequence 196137, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 196137
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_19137C.1.pep
US-10-424-599-196137

Query Match          45.5%; Score 46; DB 12; Length 73;
```

```

Best Local Similarity 50.0%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 VLPPRRKMKGLFSQAKIS 18
      : |||: |||:
Db      12 IAPPKKKKKGESPLAKIT 29

```

```

RESULT 17
US-10-424-599-265499
; Sequence 265499, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 265499
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(158)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81766C.1.pep
US-10-424-599-265499

```

```

Query Match      45.5%; Score 46; DB 13; Length 158;
Best Local Similarity 54.9%; Pred.No. 32;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY      3 P P P R X M K G L F S Q A K I S L 19
      ||| : ||| : |||
Db      92 P P P K K K G G L K E P K I V L 108

```

RESULT 18
US-10-094-749-2193
; Sequence 2193, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOPYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: NASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_73833C.1.pap
US-10-424-599-256711

Query Match 44.6%; Score 45; DB 12; Length 2834;
Best Local Similarity 47.4%; Pred. No. 9.2e+02;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISL 19
Db 1109 LMRVRKLGKGLASRKASI 1127

RESULT 21

US-10-424-599-200960
; Sequence 200960, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200960
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(151)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_23493C.1.pap
US-10-424-599-200960

Query Match 43.6%; Score 44; DB 12; Length 151;
Best Local Similarity 52.6%; Pred. No. 6.1;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISL 19
Db 76 LPPSTPTSTPLFSQKINL 94

RESULT 22

US-10-424-599-198250
; Sequence 198250, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 198250
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_21044C.1.pap
US-10-424-599-198250

Query Match 43.6%; Score 44; DB 12; Length 178;
Best Local Similarity 45.0%; Pred. No. 7.3;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20
Db 158 VRPPTKLTYYKATKPNLF 177

RESULT 23

US-10-087-684-76
; Sequence 76, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boidog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangoli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 76
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (274)..(274)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (341)..(341)
; OTHER INFORMATION: Wherein Xaa may be any naturally occurring amino acid
US-10-087-684-76

Query Match 43.6%; Score 44; DB 12; Length 408;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```
QY      3 PPRKMKGLFSQAKISLF 20
Db      186 PDPKMTQLSQAKLAVF 203

RESULT 24
US-10-218-779-76
; Sequence 76, Application US/10218779
; Publication No. US20040029222A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit
; APPLICANT: MacDougall, John
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David
; APPLICANT: Gerlach, Valerie
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook II, John
; APPLICANT: Lepley, Denise
; APPLICANT: Rieger, Daniel
; APPLICANT: Burgess, Catherine
; APPLICANT: Casman, Stacie
; APPLICANT: Spytek, Kimberly
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Patturajan, Meera
; APPLICANT: Shency, Suresh
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar
; APPLICANT: Vernet, Corine
; APPLICANT: Zehrusen, Bryan
; APPLICANT: Malyankar, Uriel
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles
; APPLICANT: Gangolli, Bsha
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-214
; CURRENT APPLICATION NUMBER: US/10/218,779
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,-926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 408
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-218-779-76

Query Match      43.6%; Score 44; DB 12; Length 408;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      3 PPRKMKGLFSQAKISLF 20
Db      186 PDPKMTQLSQAKLAVF 203

RESULT 25
US-10-425-114-54364
; Sequence 34364, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54364
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLB73030H10_FLI.pap
US-10-425-114-54364

Query Match      43.6%; Score 44; DB 12; Length 457;
Best Local Similarity 77.8%; Pred. No. 1.9e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      3 PPRKMKGL 11
Db      207 PPRKMKGL 215

RESULT 26
US-10-425-114-43063
; Sequence 43063, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 43063
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-061-G7_FLI.pap
US-10-425-114-43063

Query Match      43.6%; Score 44; DB 12; Length 490;
Best Local Similarity 77.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      3 PPRKMKGL 11
Db      240 PPRKMKGL 248

RESULT 27
US-10-104-047-3665
; Sequence 3665, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
```


;
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,020
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,876
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,895
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,884
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,894
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,971
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,964
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,882
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,899
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,893
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,900
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,901
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,892
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,915
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,019
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,970
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,972
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,916
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,373
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,875
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/049,374
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,883
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,897
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,898
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,962
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,963
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,877
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/048,878
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/073,160
; PRIOR FILING DATE: 1998-01-30

;
; PRIOR APPLICATION NUMBER: 60/073,159
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,165
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/073,164
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: 60/085,925
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,921
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-933-767-855

Query Match 43.1%; Score 43.5; DB 10; Length 97;
Best Local Similarity 47.4%; Pred. No. 46;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPRRKMKGLFSQAKISLF 20
:|||||:|:|:|:
Db 52 VPPFLKMPGVLA-IELSIF 69

Search completed: May 11, 2004, 13:44:51
Job time : 35.25 secs